A Universal Model of Diagnostic Reasoning
Pat Croskerry, MD, PhD

Abstract
Clinical judgment is a critical aspect of physician performance in medicine. It is essential in the formulation of a diagnosis and key to the effective and safe management of patients. Yet, the overall diagnostic error rate remains unacceptably high. In more than four decades of research, a variety of approaches have been taken, but a consensus approach toward diagnostic decision making has not emerged. In the last 20 years, important gains have been made in psychological research on the effort to augment and improve the diagnostic performance of clinicians.

Improving diagnostic reasoning would seem to be an important goal for the safety of patients; however, a major impediment has been the variety of approaches that have been taken toward understanding the clinical reasoning that underlies the diagnostic process. These cluster into two main groups (see List 1), following the historical division into intuitive or analytical approaches toward thinking, reasoning, and deciding. The various approaches that have been taken toward decision making have two implicit purposes: first, to explain the ways in which we think and, second, to generate a practical approach to decision making that has important clinical utilization.

The intuitive approach leans heavily on the experience of the decision maker and, therefore, uses reasoning that depends on inductive logic. Experienced decision makers recognize overall patterns (gestalt effects) in the information presented and act accordingly—action is recognition primed. The experience of the decision maker will determine how well the information presented is interpreted as the decision maker seeks to make sense of the overall gestalt. Typically, such decisions are made under uncertainty; they employ heuristics or mental shortcuts, and they may be made quickly using thin-slice sampling (ie, relying on instinctive first impressions).

As we rarely have all of the information necessary to make an informed decision, such “rational” decisions have bounds or limitations, but we do the best we can under the circumstances. In recent years, the intuitive approach has also come to incorporate elements of evolutionary psychology—the view that some of our thinking is driven by cognitive modules that are hardwired in the Darwinian sense. Also, there is accumulating interest in the role of preattentive, or unconscious, mental processes—the view that perceptual analysis can effortlessly occur without deliberate intention or awareness and lead to judgment and action.

The analytical approach, in contrast, takes place under more ideal conditions, where there are fewer boundaries and greater availability of resources, resulting in less uncertainty; decisions made under these circumstances approach normative reasoning and rationality more closely. If all the relevant variables and the parameters of test performance are known, then one can use the Bayesian method to calculate fairly exact probabilities of the likelihood of a particular disease. The analytic reasoning mode is classically Popperian, with hypothesis testing and deductive reasoning; it is analytical, involves critical thinking, and is logically sound.

Arborization, or multiple branching, is
List 1

Comparison of Intuitive and Analytical Approaches to Decision Making

Intuitive
- Experiential-inductive
- Bounded rationality
- Heuristic
- Gestalt effect/pattern recognition
- Modular (hard-wired) responsivity
- Recognition-pruned/thin slicing
- Unconscious thinking theory

Analytical
- Hypothetico-deductive
- Unbounded rationality
- Normative reasoning
- Robust decision making
- Acquired, critical, logical thought
- Multiple branching, arborization
- Deliberate, purposeful thinking

an algorithmic approach using a series of unambiguous branching points and is particularly useful for delegated decision making. Essentially, it is analytic decision making by proxy, the branching points having been researched and refined by experts in the field. The exhaustion strategy involves first collecting all possible relevant data and then searching through the data for a diagnosis. It characterizes the approach of novices, but it may also be employed by experts. When diagnoses are rare and esoteric, as well as under conditions of sleep deprivation and fatigue, robust decision making is made more analytical than intuitive. It adopts a systematic approach to remove uncertainty within the resources available to make safe and effective decisions.

Cognitive continuum theory proposes that there is not a dichotomy but a continuum between intuitive and analytical approaches.

These various approaches have their origins in the fields of mathematics, philosophy, and, predominantly, psychology, and medical decision-making strategies have generally borrowed from them. However, after decades of research activity in medical decision making, no unifying approach has emerged and, correspondingly, it has not been possible to teach a consistent approach toward medical decision making.

I propose a model here that brings together recent developments in cognitive psychological theory, these varied approaches to medical decision theory, and the realities of clinical practice.

System 1 and System 2 Processes

Description
Two fundamental approaches to reasoning, intuitive and analytical, have been formally recognized during the last 20 years. This dichotomy has since gathered momentum in the psychology literature and is now widely recognized as dual process theory, or System 1 and System 2 reasoning.

The main characteristics of the two systems are listed in Table 1. They have recently been further elaborated and clustered by Evans.

System 1 is an intuitive approach that proves effective much of the time. Importantly, it is highly context-bound, with the potential for ambient conditions to exert a powerful influence. In forming their early diagnostic impressions, physicians may be consciously or subconsciously influenced by a variety of factors, including patient characteristics (appearance, demeanor, degree of discomfort, communication issues, past experience with the patient), characteristics of the illness (acuity, severity, past experience with the presenting complaints), immediate issues in the medical environment (other patients’ needs, workload, priority setting, interruptions, distractions), resource issues (availability of specific tests, procedures, consultants, hospital beds), overarching issues (professional, ethical, medicolegal), and others.

Morbidity and mortality rounds often do not take account of these contextual and ambient conditions when cases are being reviewed; regrettably, this is also a feature of medicolegal investigations.

System 1 is characterized by heuristics and other mental shortcuts. For example, most physicians would have little difficulty recognizing the characteristic distribution and appearance of herpes zoster, or the combination of signs and symptoms of an acute myocardial infarct. Many diagnostic decisions in medicine are based on this type of pattern recognition, which is strongly related to how fully manifested the disease is (i.e., how characteristic and pathognomonic the presentation is for a particular illness). The system is fast, frugal, requires little effort, and frequently gets the right answer. But occasionally it fails, sometimes catastrophically. Predictably, it misses the patient who presents atypically, or when the pattern is mistaken for something else. In a major study of acute coronary syndrome, for example, the diagnostic error rate increased 10-fold when patients presented without the cardinal symptom of chest pain.

System 2, in contrast, is engaged when the patient’s signs and symptoms are not readily recognized as belonging to a specific illness category, or do not follow a particular script. For a patient presenting with a global headache, for example, there are a variety of diagnostic possibilities: muscle tension headache,

<table>
<thead>
<tr>
<th>Cognitive style</th>
<th>Type 1 Heuristic, intuitive</th>
<th>Type 2 Systematic, analytical</th>
</tr>
</thead>
<tbody>
<tr>
<td>Computational principle</td>
<td>Associative</td>
<td>Rule based</td>
</tr>
<tr>
<td>Responsiveness</td>
<td>Passive</td>
<td>Active</td>
</tr>
<tr>
<td>Capacity</td>
<td>High</td>
<td>Limited</td>
</tr>
<tr>
<td>Cognitive awareness/control</td>
<td>Low</td>
<td>High</td>
</tr>
<tr>
<td>Automaticity</td>
<td>High</td>
<td>Low</td>
</tr>
<tr>
<td>Rate</td>
<td>Fast</td>
<td>Slow</td>
</tr>
<tr>
<td>Reliability</td>
<td>Low</td>
<td>High</td>
</tr>
<tr>
<td>Errors</td>
<td>Relatively common</td>
<td>Rare</td>
</tr>
<tr>
<td>Effort</td>
<td>Low</td>
<td>High</td>
</tr>
<tr>
<td>Emotional attachment</td>
<td>High</td>
<td>Low</td>
</tr>
<tr>
<td>Scientific rigor</td>
<td>Low</td>
<td>High</td>
</tr>
</tbody>
</table>

rebound headache, migraine, subarachnoid hemorrhage, meningitis and others. The degree of pathognomicity is low, and uncertainty is correspondingly high; the various possibilities must now be teased out from each other in a systematic search. A System 2 approach is required—it is analytical, slow, and resource intensive, but more likely to get the correct diagnosis than would System 1.

Origins
The two systems have important and differing origins. System 1 is a passive, reflexive set of systems that may be triggered by context, images, emotions, and older parts of the brain—modules that evolved to cope with specific survival needs of our ancient evolutionary past. It is capable of parallel processing, and it is responsive to more than one feature at a time. Most System 1 responses seem to be evolutionarily adaptive, but they may not be instrumentally rational in modern contexts, leading to potential mismatches between our cognitive capabilities and prevailing environmental circumstances. It is this inherent vulnerability of intuitive thinking, and the use of heuristics that goes with it, that account for much of the error in System 1. Although System 1 reflects the innate responsivity of the brain, repetitive processing by System 2 can eventually lead to a System 1 response. For example, the first time a medical student sees a shingles rash, it will not be meaningful, but with repeat presentations the formulation of the diagnosis will eventually become reflexive.

In contrast, System 2 is the logical, rational software of the brain and only processes one channel at a time. It requires conscious activation. It is a linear system that is built through learning—the nurture part of our reasoning faculties. It becomes increasingly competent as we mature, socialize, and go through formal education. It is refined by training in critical thinking and logical reasoning.

Operating Characteristics of the Model
The unmodified process
The first step in the diagnostic reasoning process is the presentation of the patient’s symptoms and signs to the decision maker (see Figure 1). This is usually through direct contact between physician and patient, but it may take a less proximate form in which the patient’s signs, symptoms, and results of investigations are relayed to a physician through an intermediary, such as a junior house physician to an attending staff member, or a family physician to a consultant. In the learning context, virtual patients or written descriptions of a patient’s illness may be used to teach about the diagnostic process. In my experience, some fidelity of information is often lost in these second-hand accounts because of the intrusion of the first observer’s thinking and interpretation biases, as well as a loss of context and ambient influences. A similar concern holds for the ecological validity of research in medical decision making that is removed from real clinical practice.

If salient features of the presentation are initially recognized, System 1 processes engage immediately and automatically. Thus, recognized visual presentations of illness or injury (e.g., dermatological conditions, dislocations, fractures, stigmata of particular diseases such as alcoholism, endocrine disorders,
cardiovascular disorders) or recognized combinations of salient symptoms or findings (syndromes, toxidromes, illness scripts, compiled experiences) will trigger pattern-recognition types of responses in System 1. Importantly, this process is reflexive and unconscious—no deliberate thinking effort is involved. These responses can only occur through prior System 2 learning; the more pathognomonic the presentation, the stronger the response to it. Studies of expert decision making strongly support the success of this pattern-recognition approach.28,29

This recognition-primed processing forms the basis of a variety of approaches to decision making, as described above,13,31,32 and broader views of the control of conscious processes.16,33 As noted, these various approaches fall into two general categories: the intuitive approaches shown in List 1 are predominantly based on System 1, whereas the analytical ones are based on System 2. Approaches such as robust decision making20 and cognitive continuum theory34 are a combination of the two systems. Fuzzy trace theory is also a dual-process-type model of reasoning but with an emphasis on memory and perception of risk.34 It distinguished two forms of representation: verbatim and gist. The latter has many characteristics of System 1 and appears similarly vulnerable to inconsistencies in reasoning and irrational biases.

In addition to the pattern-recognition response, other System 1 responses may be generated simultaneously, in parallel to that response. For example, often the first responses that physicians have toward patients involve their (the physicians’) feelings, and these may vary in both intensity and polarity. Physicians may have positive feelings toward some patients but negative feelings toward others,5,36 and often they may be unaware that these preconscious affective dispositions can play a significant role in decision making.16,37 Other System 1 responses can be triggered simultaneously with the pattern-recognition response, such as heuristics (mental shortcuts, rules of thumb), intuitions, and others. Some known determinants of System 123 are shown in the model. System 1 decision making has been popularized in the book Blink38 as the rapid cognitive style mentioned earlier called thin slicing,13 although this approach may, in certain circumstances, prove perilous in medicine.39,40

If the presentation is not recognized, or if it is unduly ambiguous or there is uncertainty, System 2 processes engage instead. Now the system is an analytic one, attempting to make sense of the presentation through objective and systematic examination of the data, and by applying accepted rules of reasoning and logic. It is a linear processing system, slower than System 1, more costly in terms of investment of resources, relatively free of affective influences, but considerably less prone to error. Some of the factors that influence System 2 reasoning are shown in the model.23 If there are no subsequent modifications of System 1 or System 2 processing, their individual or blended outputs determine the calibration of response and the eventual veracity of the diagnosis.

Modifiers of the process

The model has several mechanisms for modifying its output. First, System 1 and System 2 may interact with each other so that the final output is a synthesis of the two.23 For example, the patient’s initial presentation might trigger a System 1 response in the decision maker that subsequently sets up a System 2 analytical approach. The monitoring capacity of System 2 over System 1 allows it to reject the latter by applying a rational override. Thus, while the first look at a rash might trigger a shingles diagnosis, if there are aberrant or atypical features (it crosses the midline or doesn’t follow a dermatome distribution), System 2 can override and force a reassessment. Importantly, inattentiveness, distraction, fatigue, and cognitive indolence may all diminish System 2 surveillance and allow System 1 more latitude than it deserves. With fatigue and sleep deprivation, for example, the diagnostic error rate can increase fivefold.41 The monitoring capacity of System 2 depends on deliberate mental effort and works best when the decision maker is well rested, well slept, free from distraction, and focused on the task at hand. Metacognition, the ability to step back and reflect on what is going on in a clinical situation, is essentially System 2 monitoring in action, and may save a critical miss from occurring. Quirk42 defines metacognition as the act of “thinking about one’s own and another’s thinking and feeling.” It forces a monitoring step similar to the one that occurs when clinical decision support tools are used.

Second, System 1 may override otherwise sound reasoning developed by System 2. Consider, for example, a physician who may have attended a teaching session on a clinical decision rule for determining the pretest probability assessment of pulmonary embolus, or who has read about it in a journal or discussed it with a colleague and decided that this is the most rational and optimal approach to follow in this particular clinical situation. Such decision rules are based on aggregate data and developed objectively through System 2 reasoning and investigation in the cold light of day, as in the arborization, multiple-branching approach. However, when the physician is faced with a particular patient in a real-life situation, the physician may choose to override the decision rule and follow his or her intuitive feelings. Occasionally, there may be virtue in this; “situational appreciation” in some circumstances may be important for determining what is appropriate in a particular situation,43 but as an overall strategy it may prove irrational. These override decisions are not uncommon in medicine and may underlie, in part, the difficulties in acceptance and incorporation of clinical decision rules, referred to as knowledge uptake, transfer, or translation. Essentially, these inconsistencies, quirks, self-deceptions, and variances in individual decision making represent departures from a rational approach; they occur for historical, habitual, emotional, situational, and a variety of other reasons. Thus, even though well-developed clinical decision guidelines may be shown to consistently outperform the decision-making capabilities of the individual physician, there may still persist an irrational belief in some individuals that they know best and can always do better for the patient. This overconfidence is a major source of diagnostic error.44,45

In presenting the model, I am not suggesting that all medical reasoning and decision making fall neatly into one or the other system. As has been noted, instead of a discrete separation of the two systems, a cognitive continuum with oscillation occurring between System 1 and 2 has been proposed, resulting in
varying degrees of efficiency and accuracy in judgment. When conflict occurs between competing goals of the two systems, it might be prudent and safer for the patient if the clinician applied a System 2 override of System 1. In recent reviews, the oversimplification that has occurred in dual-process approaches, and the inherent difficulties in accommodating the multiple and heterogeneous attributes of Systems 1 and 2, have been noted. Evans suggests that it might be more appropriate to talk about type 1 and type 2 processes, and Stanovich has recently proposed a subdivision of System 2 into algorithmic and reflective levels. The algorithmic mode engages the intellectual abilities and cognitive strategies of the decision maker, whereas the higher-level reflective mode involves beliefs, overall goals, and general knowledge.

Models of cognitive reasoning are relatively slow works in progress. Although the model proposed here may appear to simplify the complex processes at work in reasoning, decision making, and judgment, it nevertheless provides a basic framework for medical decision making within a sound theoretical structure, incorporating the disparate and diverse approaches that have been observed historically.

Clinical Relevance Versus Scientific Rigor of the Two Systems

The dual-processing theory can be used to bridge the current division of approaches toward clinical reasoning and decision making. Orthodox medical decision theorists have historically emphasized a scientifically rigorous approach toward decision making that is typically based on statistical and mathematical models. An inherent and prevailing assumption, concerning what are predominantly System 2 approaches, is that they can be employed by well-rested, well-slept decision makers under ideal conditions in which there are no distractions or untoward intrusion of affect and all the required data are available. However, as Reason notes, the cognitive reality often departs from this formalized ideal.

In many medical settings, decision makers function under suboptimal conditions. They may be hurried, distracted, fatigued, sleep deprived, and limited by resource constraints. Shortcuts and heuristic reasoning may come into play under conditions of cognitive busyness, overload, noisy signals, fatigue, and resource limitations. In many medical settings, workload is dynamic, often varying unpredictably, and providers must select strategies to maintain throughput of patients. One obvious strategy is to attenuate workload by using heuristics and shortcuts that achieve speed and frugality of cognitive effort—what has been referred to as the cognitive miser assumption. However, these strategies, which characterize System 1 approaches, are known to be vulnerable to a variety of cognitive and affective biases. Further, the area of the brain believed to be the neuroanatomical substrate for System 2 reasoning—the anterior cingulate cortex, prefrontal cortex, and medial aspect of the temporal lobe—is the same area that suffers neurocognitive compromise through sleep deprivation. Thus, the combined effects that occur under adverse working conditions (i.e., the increased use of heuristics, together with the functional compromise of System 2) result in decrements in clinical performance, especially those aspects of performance associated with decision making.

Finally, many clinical situations are often characterized by too many variables or unknowns, too many ethical and financial restrictions, or too many other resource limitations to ever allow a simple quantitative approach to guide a particular clinical decision, and actuarial models simply cannot be applied in many clinical situations. This is the clinical reality that medical decision makers face daily.

Thirty years ago, this dilemma was presciently recognized by Elstein as a clinical-statistical polarization. The prevailing perception among medical decision makers at that time was that there was no scientific worth without quantification and statistics (i.e., the System 2 approach), a view that is little changed today: “The broader community of medical decision making researchers,” stated Hamm, “has not embraced the topic of heuristics and biases approach with sustained enthusiasm.” Failure of the theory of heuristic strategies (System 1 approach) has been attributed to its weak predictive power, its inability to describe the judgment process in sufficient detail or to explain individual differences, and its failure to assist physicians in improving their decision making. So, the paradox remains: an important feature of clinicians’ decision making is apparently disqualified from study by those who research the field of medical decision making. This situation does little to alleviate the discomfort of physicians whose “dilemma,” stated Hammond, “lies in the rivalry between intuition and analysis. Intuition offers an immediate if risky judgment; analysis, though safer, takes longer—if it can be done at all.” Periodically, articles are published describing selected heuristics and biases in clinical reasoning with caveats about their pitfalls and failures, but the overall breadth of the problem is not addressed. There are over 50 known cognitive biases, many with evident influence in clinical decision making, and a variety of affective biases. To date, there has been little research on the role these biases play in real clinical decision making.

It is important to resolve this issue for several reasons. First, clinical decision making is a critical aspect of clinical performance. In fact, it is difficult to imagine anything of greater importance or relevance to patient outcomes and to patient safety. Yet, it seems that insufficient emphasis is being placed on core aspects of decision making that are integral to clinical practice. Second, the failure to conduct clinical research in this area has led to a general pessimism about developing strategies to overcome the undesirable effects of heuristics and biases—that is, to develop cognitive and affective de-biasing strategies. Again, as Elstein observed, “The more it is insisted that a clinical situation cannot be analyzed in terms of risks and likelihoods, estimated however roughly, the more investigation in these terms is discouraged.” Third, clinicians in training and those already in practice need a comprehensive approach to clinical decision making that facilitates their understanding of this complex process and allows them to gain insight and understanding into their own decision making. For the safety of patients, the imperative to think critically, reason, decide, and diagnose well always remains.
References


---

Did You Know?

In 1981, surgeons at the UCSF School of Medicine performed the first successful surgery on a fetus still in the womb.

For other important milestones in medical knowledge and practice credited to academic medical centers, visit the “Discoveries and Innovations in Patient Care and Research Database” at (www.aamc.org/innovations).