



PROJECT MUSE®

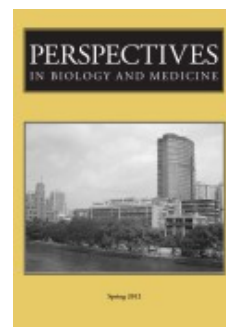
Clinical Expertise and the Limits of Explicit Knowledge

Daniel Luchins

Perspectives in Biology and Medicine, Volume 55, Number 2, Spring 2012, pp. 283-290 (Article)

Published by The Johns Hopkins University Press

DOI: [10.1353/pbm.2012.0023](https://doi.org/10.1353/pbm.2012.0023)



➔ For additional information about this article

<http://muse.jhu.edu/journals/pbm/summary/v055/55.2.luchins.html>

CLINICAL EXPERTISE AND THE LIMITS OF EXPLICIT KNOWLEDGE

DANIEL LUCHINS

ABSTRACT This article questions the view that medical decision-making can be reduced to a series of explicit rules, adherence to which will necessarily improve outcomes. Instead, it attempts to rehabilitate the concepts of clinical expertise and clinical experience, arguing that medicine, like other areas of expertise, depends on forms of implicit knowledge that can only be acquired through years of experience. Recent research on “fast and frugal” heuristics in medical decision-making suggest that statistical techniques are not necessarily superior to clinician judgment. Since clinical decisions are made on individual patients within the constraint of limited information, they must rest on clinical expertise and not clinical rules.

CONTEMPORARY MEDICINE is subject to forces determined to make public and explicit the criteria used in clinical decisions. Within our profession, the promotion of evidence-based medicine has elevated findings from meta-analysis of controlled clinical trials into a gold standard for practice. In this context, multiple professional organizations have established guidelines for appropriate care. Outside our profession, the well-publicized variations in regional medical practice patterns have helped undermine confidence in physicians’ clinical judgment, while at the same time strengthening efforts by the bureaucracies

Department of Psychiatry, University of Illinois at Chicago, and Mental Health Research, Jesse Brown Veteran Affairs Medical Center, Chicago.

Correspondence: Mental Health Research, Jesse Brown Veteran Affairs Medical Center, 820 South Damen Avenue, Chicago, IL 60612.

E-mail: daniel.luchins@va.gov.

The author would like to thank Bill Salas and John Monahan for their helpful suggestions.

Perspectives in Biology and Medicine, volume 55, number 2 (spring 2012):283–90

© 2012 by The Johns Hopkins University Press

that oversee and reimburse our industry to apply detailed regulations that would promote consistency across clinicians and regions. Even in the clinical encounter, widespread internet access to medical information places physicians in the position of having to justify their decisions to increasingly informed and critical consumers.

Doubtless, these developments have their benefits and are probably inevitable in liberal democracies that question traditional sources of authority. In any case, one could argue that medical practitioners are better off appealing to science as the ultimate arbitrator than to any of the alternatives. Yet on the other hand, medicine, despite its allegiance to science, has also always considered itself an art, something that cannot be reduced to principles. Aristotle, the son of a physician, included medicine in this realm of practical knowledge: “Matters of conduct have nothing fixed or invariant about them, any more than matters of health. . . . Just as in the arts of Medicine and Helmsmanship, agents have to consider what is suited to the occasion” (Toulmin, 2001, p. 109). Even in the 1980s, commentators could still claim that the “two cultures of biomedicine, science and practice” were neither identical nor hierarchical, that “scientific norms concern knowledge . . . not practical outcomes,” and that practice must be guided by what is particular and local (Greer 1987). Today, such claims would be viewed with suspicion, an excuse for not following generally accepted guidelines. Likewise the physicians’ appeal to “clinical expertise” based on “clinical experience” would be dismissed as “a prestigious synonym for anecdotal evidence when anecdotes are told by somebody with a professional degree and a license to practice a healing art” (Grove and Meehl 1996, p. 302).

ACTUARIAL VERSUS CLINICAL DECISION-MAKING

Skepticism regarding the clinician’s claim to expertise is in keeping with a scientific *weltanschauung* that views human decision-making as impaired (Tversky and Kahneman 1974). Kahneman received the Nobel Prize in 2002 for research supporting just such a view. Such skepticism is also consistent with an extensive literature, going back a half century that specifically examined the judgment of clinical experts and found it wanting. This literature has had demonstrated that “actuarial” decision-making procedures—those relying on even simple statistical methods—consistently outperform clinical experts. For example, when pathologists coded biopsies of patients with Hodgkin’s disease and then made an overall rating of the tumors’ severity, the correlation of rating to survival time was virtually zero. However, if the variables that the doctors identified were used in a multiple regression equation, they could successfully predict survival time (Einhorn 1972). This finding is an example of what had been described as the “Goldberg paradox”: data from the same criterion used by recognized experts to make decisions produces more accurate decisions when used in some formal process than when used by the experts themselves. These studies would seem to

support the view that patients would be better served if clinicians relied on explicit actuarial techniques and not on their so-called clinical expertise.

Yet efforts over the last few decades to integrate actuarial techniques into clinical practice have generally failed. In part, this is because while these actuarial versus clinical studies ostensibly involved clinical judgment, they were not carried out in a clinical context but in the laboratory. They examined single, discrete decisions, involving a small numbers of options. In effect, the decision points were extracted from the ongoing clinical context, in which the greater problem for the clinician is knowing when and how to frame a decision with a small number of options (Wears and Berg 2005).

But even in clinical situations in which a decision point could be clearly defined, actuarial techniques often have not proven useful. One such intervention tried to improve family practice residents' decision to admit patients with suspected myocardial infarctions (MIs) to a CCU (Green and Mehr 1997). Before the study began, 90% of suspected patients were admitted, including a high rate of false positives. Using retrospective data from that setting, the authors tested different decision-support tools and found that high scores on the Heart Disease Predictive Instrument (HDPI), based on the presence or absence of seven historical and EKG findings, was highly predictive of MIs. Encouraged by these findings, they initiated an ABAB reversal design: over the course of one year, they distributed and took away a laminated card containing the HDPI and instructed the residents to use them along with a pocket calculator to make more accurate predictions. To their surprise, the authors found that even prior to receiving the cards (but after hearing about them), and also after they were taken away, the residents' decisions had improved with sensitivity and specificity comparable to what could be achieved if they had actually used the HDPI. What the residents had learned was not the complicated mathematical model of the HDPI, but the clinical "cues" to focus upon.

Green and Mehr's findings are consistent with a growing literature showing that human decision-making does not rely on the integration of multiple cues, but on the sequential processing of a small number of cues using "fast and frugal" heuristics (Gigerenzer and Kurzenhauser 2005). Furthermore, in conditions of uncertainty or limited information, these heuristics are equal to or at times even superior in accuracy to regression-based models. Several such heuristics ("take the best," "simple tally") could have explained the residents' performance. One possibility with extremely good predictive value would be a simple "two-step rule": if there is an ST abnormality, admit to the CCU; if no ST change but a chief complaint of chest pain and any other abnormality, also admit; otherwise don't. Which, if any, of these heuristics was actually used by the residents is unknown, as they were never asked. But a study of physicians' decisions to prescribe cholesterol-lowering drugs that queried the clinician about the number and order of cues used suggests that these processes are often not conscious (Dhani and Harries 2001).

The mathematical modeling that underlies the HDPI or comparable scales may be better than any heuristic in accommodating the available data, while not necessarily being as good at predicting future results. This is because such modeling “over fits” the data, or too closely models previous data at the expense of predicting future data (Gigerenzer and Kurzenhauser 2005). For example, starting with a set of patients’ seven HDPI variables and whether they did or did not have an MI, it would be easy to come up with a complicated mathematical model that better fit the data than the “two-step rule.” The model would better fit the previous data, yet it would quite likely lack the robustness that allows it to generalize to new data sets and thus not be useful in predicting the outcome in the next case.

PREDICTING VERSUS TREATING

Furthermore, clinicians are not meteorologists: they don’t want to simply predict what will happen but rather want to intervene to change the outcome. To do this, they need to make inferences that tie together cause and effect in a way that can be used to guide interventions. Being able to predict the life expectancy of a person with certain physical findings is not as clinically important as determining what is causing these findings and therefore how it should be treated. With statistical techniques (unless one has near limitless and perfect data), casual factors may remain obscure.

In support of this idea, Einhorn and Hogarth (1982) provide an amusing example. A statistically but not reproductively sophisticated tribe is undergoing an alarming decrease in its birth rate and needs to determine the cause of birth. Sexual intercourse is hypothesized to be the cause, but many alternatives (such as sitting under a palm tree in the full moon holding hands) cannot be ruled out. The tribe randomly allocates 100 women to an “intercourse” condition, of whom 20 become pregnant, and 100 to “non-intercourse,” of whom five become pregnant (because of “measurement error” due to poor memory, lying, etc.). The experiment provides only a modest ($r = 0.34$) correlation between intercourse and pregnancy, and the unsupported theory is therefore discarded in favor of the palm tree hypothesis.

Statistical reasoning and causal reasoning are quite separate tasks. Humans, including experts, have great difficulty with statistical reasoning, in part because we want to impose a causal order on situations that are random (Tversky and Kahneman 1974). For example, if subjects are rewarded for correct predictions when a red light is randomly presented 60% of the time and a green light 40% of the time, they could maximize their performance by always choosing red (60% correct). Instead, because they want to correctly predict a random event (as opposed to maximizing their odds), they will display “probability matching” by choosing red 60% of the time and green 40% of the time (52% correct). In such situations, our need to assign causality and to believe that we can predict random events

not only leads to poor predictions, but to a variety of cognitive distortions, including superstitious beliefs and the “illusion of control” (Einhorn and Hogarth 1982). But that is only true if the phenomenon is truly random. If there are causal relationships on which our interventions could have an impact, adherence to a statistical approach maximizes short-term predictive accuracy at the expense of future knowledge and cure. The history of medicine would suggest we have been well served by the clinicians’ search for casual models and reluctance to accept stochastic ones.

Even when casual relationships are understood, salient factors known, and the fast and frugal heuristics used by clinicians can be identified, turning these clinical tools into explicit rules can be misguided. There is a growing body of evidence documenting the pitfalls of converting clinical guidelines into quality measures (Walters et al. 2004). Clinicians are concerned with the individual patient, not representative populations, so clinical expertise will remain essential in determining when and with whom adherence to any particular decision procedure is appropriate. At its most basic, this is the so-called “broken leg countervailing effect” identified by Grove and Meehl (1996). A decision algorithm can predict with a high level of certainty that a particular horse will win today’s race, but the race track habitué knows this can’t be true: the horse broke its leg this morning. In the context of caring for the individual patient, this means being able to use clinical expertise to determine which if any of the characteristics of this particular individual at this particular moment would require modification of the explicit rules. Knowing what to look for to determine which of the many variables would or would not require such modification is similar to knowing what variables are relevant in viewing a biopsy or X-ray, although it may require even greater experience.

EXPLICIT VERSUS IMPLICIT KNOWLEDGE

But how does the expert decide which variables are relevant and which are irrelevant? How do experts know what to look for? By using the knowledge that depends on their specialized training and their years of experience, their clinical expertise. This is more than facility in applying the rules by which medicine should be practiced: like other forms of expertise, it is based on implicit knowledge that can only be acquired by experience and cannot be replaced by knowledge that is public and explicit. That is not to say that this knowledge is occult. It is shared by other experts and is transmitted (in part) along with the field’s explicit knowledge during one’s clinical training. It is not by reading the instructions, but only by actually following them that this knowledge can be acquired.

Psychologists distinguish between the knowledge that one can “declare” or be explicit about, and “non-declarative” or implicit knowledge (Mathews et al. 1989). Explicit knowledge depends on conscious processes, “attempts to form a mental representation, searching memory for knowledge of analogous systems

and attempts to build and test mental models” (p.1083). On the other hand, implicit knowledge is “automatic” and “nonconscious” (p.1083): one can correctly perform a task without knowing that one knows how. A familiar example of implicit knowledge is a young child’s ability to correctly use the rules of grammar without being able to describe them, and long before they are taught in school. Psychologists have modeled this phenomenon by studying the acquisition of artificial grammars (strings of letters that must follow certain rules; for example, *p* is always followed by *c*) and found little relationship between a subject’s ability to generate correct answers and correctly describe the grammar’s rules. In fact, early on, attempting to explicitly discover the rules of the artificial grammar can interfere with performance. On the other hand, when the artificial grammar has few rules or when these are readily apparent (for example, letters always follow the order of the alphabet), the results are reversed, and using knowledge of the explicit rules is favored. In general, implicit learning and the use of implicit knowledge seems to be favored when tasks involve multiple factors with no apparent way to differentiate the relevant from the irrelevant. That being said, in real-world situations, the teaching of complex tasks, including medicine, involves the appropriate mix of both forms of learning (Mayer 2004).

Since a subject’s ability to differentiate correct from incorrect letter combinations when learning an artificial grammar does not rely on formulating rules (explicit knowledge), psychologists believe that subjects instead compare the novel string with examples they have previously experienced and against which they can test their fast and frugal decision heuristics. With extensive experience, subjects are able to describe and put into an explicit form some of the rules that might underlie the grammar. However, when subjects who have not had actual experience with the task use these rules, they perform more poorly than the original subjects (Mathews 1989). This suggests that some of the knowledge of the artificial grammar is experiential and cannot be fully transmitted by explicit rules. Similarly, decades of research on clinical expertise suggests it also is experiential and cannot be fully transmitted by acquiring rules or even specific strategies, but rather by the opportunity to practice and receive feedback from multiple examples (Norman 2005).

THE NATURE OF EXPERTISE

In general, expertise is dependent on such implicit knowledge, built up of exemplars that must be acquired through experience. This explains how the blindfolded chess master, unlike the novice, is able to code the board in the appropriate way and therefore “see” the correct move (DeGroot 1965). It is not simply a matter of memorizing the positions of each piece on the board, because if the pieces are placed randomly on the board, experts do little better than novices. But when the positions are drawn from actual play, the expert far outperforms

the novice by coding the positions, using the 50,000 to 100,000 functionally significant patterns (“chunks”) that the expert has learned to recognize in the decade that it takes to become a master (Ross 2006). Since they rely on such implicit knowledge, chess experts may not be fully explicit about what they are doing. However, research by psychologists has led to a better understanding of how expert chess players know what to look for, and detailed computer models of pattern recognition can account for many of the findings with chess experts.

However, the chess-playing computers such as Deep Blue that can now outplay even grand masters do not use these techniques. Instead, the computers rely on explicit decision-making algorithms to evaluate millions of possible moves on defined criteria (such as gains in center control or protecting the king). Grand masters use “progressive deepening” to follow up on a few potentially good moves. As described by Klein (2001): “They evaluate each move separately. . . . They tried to see the overall lines of play that were created or blocked. They tried to understand the causal dynamics of the position. . . . Causal reasoning replaced probabilistic reasoning” (p. 115). Over time, computers will get better at playing chess, not by more accurately mimicking how grandmasters think, but by doing what our brains cannot. In the same vein, despite the increasing success of the computer programs, future generations of grandmasters will try to improve their play, not by mimicking computers, but by playing against them.

There is a lesson here for medicine, but it is not that clinical expertise will soon be matched and bettered by computers using decision-making algorithms and regression equations. Chess is played on 64 squares with 32 pieces that move according to fixed rules. Although incredibly complicated, all the relevant factors that affect outcome are known and can be calculated. The practice of medicine, like all complex real-world tasks, involves making decisions about individuals without full understanding of the rules and with insufficient information even about a particular case, a situation where implicit knowledge is favored. The lesson is that despite the computer’s success, chess masters are not trying to mimic its algorithms but are using them to increase their experience and enhance their expertise. Chess-playing computers provide an untiring opponent, whose skills can be adjusted to teach the novice and challenge the expert. This has made it easier to acquire the thousands of hours of experience that makes one a master, leading to a generation of child prodigies and improvement in the level of play even at international tournaments.

Although decision-making tools that rely on explicit rules may assist the clinician, even the most computationally sophisticated application of explicit rules will not insure good clinical care or identify good clinicians. In this regard the Joint Commission on Accreditation of Healthcare Organizations’ list of 405 hospitals that achieved the highest levels of adherence to clinical protocols is illustrative. The list is overrepresented by small rural facilities and VA hospitals (which provide salary bonuses to staff for adherence to these protocols), but in-

cludes none of America's premier academic institutions nor any of the 17 on the *US News and World Report* "Best Hospitals Honor Role" (Sack 2011). Since the ability to judge clinical expertise, like clinical expertise itself involves more than following explicit rules, one suspects that even those who have promoted JCAHO's rating would actually rely on the judgment of medical professionals, not such a list, when choosing where they or a loved one will receive care.

REFERENCES

- DeGroot, A. D. 1965. *Thought and choice in chess*. The Hague: Mouton.
- Dhami, M. K., and C. Harries. 2001. Fast and frugal versus regression models of human judgment. *Thinking Reasoning* 7:5–27.
- Einhorn, H. J. 1972. Expert measurement and mechanical combination. *Organ Behav Hum Perform*:86–106.
- Einhorn, H. J., and R. M. Hogarth. 1982. Prediction, diagnosis and causal thinking in forecasting. *J Forecasting* 1:23–26.
- Gigerenzer, G., and S. Kurzenhauser. 2005. Fast and frugal heuristics in medical decision making. In *Science and medicine in dialogue*, ed. R. Bibace et al., 3–15. Westport, CT: Praeger.
- Green, L., and D. R. Mehr. 1997. What alters physicians' decisions to admit to the coronary care unit? *J Fam Pract* 45:219–26.
- Greer, A. L. 1987. The two cultures of Biomedicine: Can there be consensus? *JAMA* 258: 2739–40.
- Grove, W., and P. Meehl. 1996. Comparative efficacy of informal (subjective impressionistic) and formal (mechanical, algorithmic) prediction procedures: The clinical-statistical controversy. *Psychol Public Policy Law* 2:293–323.
- Klein, G. 2001. The fiction of optimization. In *Bounded rationality: The adaptive toolbox*, ed. G. Gigerenzer and R. Selten, 103–22. Cambridge: MIT Press.
- Mathews, R. C., et al. 1989. Role of implicit and explicit processes in learning from examples: A synergistic effect. *J Exp Psychol Learn Mem Cogn* 15:1083–1100.
- Mayer, R. E. 2004. Should there be a three-strike rule against pure discovery learning? The case for guided methods of instruction. *Am Psychol* 59:14–19.
- Norman, G. 2005. Research in clinical reasoning: Past history and current trends. *Med Educ* 39:418–27.
- Ross, P. E. 2006. The expert mind. *Sci Am* 295:64–71.
- Sack, K. 2011. Report finds improved performance by hospitals. *NY Times*, Sept. 15.
- Toulmin, S. 2001. *Return to reason*. Cambridge: Harvard Univ. Press.
- Tversky, A., and D. Kahneman. 1974. Judgment under uncertainty: Heuristics and biases. *Science* 185:1124–31.
- Walters, L. C., et al. 2004. Pitfalls of converting practice guidelines into quality measures. *JAMA* 291:2466–70.
- Wears, R. L., and M. Berg. 2005. Computing technology and clinical work: Still waiting for Godot. *JAMA* 293:1261–63.