Hindsight bias, outcome knowledge and adaptive learning

K Henriksen, H Kaplan

The ubiquitous nature of hindsight bias is a cause for concern for those engaged in investigations and retrospective analysis of medical error. Hindsight does not equal foresight. Investigations that are anchored to outcome knowledge run the risk of not capturing the complexities and uncertainties facing sharp end personnel and why their actions made sense at the time. Important lessons go unlearned if the exercise is simply to back track someone else’s decision landmarks. Outcome knowledge can also bias our thinking on the quality of the processes that led to the outcome. This paper examines the influence of outcome knowledge in relation to reconstructive memory and legal testimony, ways for reducing the impact of outcome knowledge, and an adaptive learning framework that places hindsight bias in a broader context of rapid updating of knowledge.

One of the major challenges facing those who are called upon to investigate adverse events, not only in medicine but in any field, is understanding how knowledge of the outcome of an adverse event influences their thinking and assessment of the event. After a horrendous mishap, questions are raised as to why those involved did not anticipate it. We tend to expect that others should know by foresight what we have learned by hindsight. The problem is that this bias is pervasive and has the potential to impact adversely on a wide variety of professional settings, especially medicine and health services research with their fondness for retrospective techniques (such as morbidity and mortality conferences, malpractice claims analysis, error reporting systems, chart reviews) for investigating adverse events that harm patients.

BEING SMART AFTER THE EVENT

Proclamations about medical error are almost always made after the event, rarely before. As noted by Reason, the most significant psychological difference between individuals who were involved in events leading up to a mishap and those who are called upon to investigate it after it has occurred is knowledge of the outcome. Investigators have the luxury in hindsight of knowing how things are going to turn out; nurses, physicians, and technicians at the sharp end do not. While most people would not expect much credit for picking a horse after it has won the race, many investigators are unaware of the influence of outcome knowledge on their perceptions and reconstructions of the event. Given the advantage of a known outcome, what would have been a bewildering array of non-convergent events becomes assimilated into a coherent causal framework for making sense out of what happened. In fact, it may be difficult to imagine it happening any other way. “Why couldn’t they see it?” is the question that is often asked. Such hindsight results in expectations by investigators that participants should have anticipated the mishap by foresight. If investigations of adverse events are to be fair and yield new knowledge, greater focus and sensitivity needs to be given to recreating the muddled web of precursory and proximal circumstances that existed for personnel at the sharp end before the mishap occurred. Dekker aptly reminds us that the point of investigations of human error is not to find where people went wrong, but to understand why their assessments and actions made sense at the time. Progressively working backwards through a causal framework with a known outcome runs the risk of delusional clarity and important lessons unlearned.

SOME EARLY RESEARCH

Convincing demonstrations of hindsight bias are well established experimentally and the phenomenon has been the subject of extensive reviews. Robust reports of it come from a variety of domains including medical diagnoses, legal rulings, financial forecasts, election returns, business outcomes, sporting events, and military campaigns.

The following case history is representative of the experimental procedure. Imagine as a practising physician that you are given a case history to read of a 37 year old bartender who developed increasing pain in his left knee which had become hot and swollen. A few days later similar symptoms develop in his left wrist and right kneec. You also receive laboratory test results, but not all of them are back. Your task is to assign probabilities to each of four given diagnoses so that the probabilities sum up to 100%. This task was given to 75 physicians divided into five groups. The first group, called the foresight group, was given the case history and simply asked to assign probability estimates to each of the four possible diagnoses. The remaining four groups were hindsight groups, each of which received the same information except for a different first sentence as follows: (1) this is a case history of Reiter’s syndrome; (2) this is a case history of post-streptococcal arthritis in an adult, (3) this is a case history of gout, and (4) this is a case history of hepatitis. As the instructions requested, all gave their independent diagnosis of the patient. The physicians exhibited evidence of hindsight bias, but the bias
was confined to the two diagnoses (post-streptococcal arthritis and hepatitis) that were assigned the lowest probability estimates by the foresight group. When compared with the foresight group, these two hindsight groups were 2–3 times more likely to make a diagnosis in accordance with the outcome knowledge expressed in the first sentence. Not all the diagnoses were equally plausible as shown by the higher probabilities given by the foresight group to Reiter’s syndrome (44%) and gout (29%). As noted by Arkes et al, physicians are well equipped with a body of knowledge to make judgment and informed decisions. Yet these highly knowledgeable physicians tilted in the direction of the outcome statement when the less probable diagnoses were presented. Similar susceptibility to hindsight bias occurs in clinicopathological conferences where less experienced physicians succumbed to the bias with both easier and more difficult cases, while more experienced physicians manifested the bias only with the easier cases.12

practical implications

What implications do these findings have for the daily practice of medicine, or any other profession where diagnosis, problem solving, or making sense of past events are central activities? The psychological conditions that exist when a physician is brought in to render a second opinion or assume care of a patient previously managed by another physician are similar to the conditions that exist when an accident investigator takes over a case. There is usually a first opinion or preliminary analysis that has already been rendered. Will the investigator or clinician be able to shed the yoke of what is already seemingly known? Will there be a greater tendency to arrive at an assessment that corroborates the first opinion than in a situation where there is no previous opinion? To preclude the adverse impact of hindsight bias and to be of optimal value, the second opinion needs to be arrived at without the stain of the first.

The potential of hindsight bias also casts suspicion on educational practices that rely heavily on case histories where the outcomes are generally known or where there is a senior, all-knowing instructor who readily imparts his or her wisdom. Hindsight does not equal foresight. Looking back down the path of another person’s journey is not the same thing as making the trip yourself. Students may gain a false sense of their own diagnostic abilities if they only retrace the moves along someone else’s decision path while experiencing the uncertainty, conflicting information, multiple possibilities, and absence of a known outcome. An unfortunate consequence of hindsight bias is that it can hamper new learning.13 If we are unable to reconstruct the intricacies of an event because of the weight that knowledge of the outcome carries (“we knew it all along”), we will not learn much from the experience.

In distinguishing between hindsight and foresight, it is instructive to examine the role each plays in morbidity and mortality and clinicopathological conferences. With the morbidity and mortality conference an adverse outcome is known and serves as the starting point as one works backwards to connect the dots and find the culprits—a technique conducive to meting out a fair measure of blame and shame. In clinicopathological conferences a discussant presents an interesting case and, after giving consideration to a number of possible diagnoses, shares with the audience the final diagnosis. At the outset, clinicopathological conferences appear to be doing the right thing by starting along a foresight path with an unknown outcome—a path characterized by uncertainty and case difficulty. Yet, before long the audience often anticipates where the discussant is heading or is simply told the diagnosis. If insufficient attention is given to portraying the uncertainty and other possibilities that exist along the foresight path, then audience members, placed in an essentially passive observer role, can also fall victim to hindsight bias and wonder why the discussant couldn’t see the obvious.12 For maximum instructional value the discussant needs to withhold his or her final diagnosis from audience members who, in turn, need to become active participants in working out the diagnosis. Careful attention needs to be given to recreating the partial information and uncertain conditions which existed along the foresight path at the time critical decisions were being made. In the absence of a known outcome, the participant’s task of identifying alternative courses of action as he/she works progressively forward is a much more challenging, engaging, and realistic exercise (for example, “Given what we know thus far, what should we do next, Dr X?”).

The deleterious effects of hindsight bias can be easily compounded by another human shortcoming known as the attribution error. Rather than giving careful consideration to the full range of situational, organizational, and system related factors that are present when adverse events happen, humans serving in a supervisory or evaluative capacity tend to make dispositional attributions and view the mishap as evidence of an inherent character flaw or defect on the part of the individual involved.15 It is only when these same personnel are themselves active agents in a mishap that they are likely to consider the role of situational and organizational variables as causal agents rather than focus on their own flawed dispositions. Our present tort system provides fertile ground for the unconstrained operation of hindsight bias and attribution error. Starting with a known outcome (that is, a patient who has been harmed), it is a fairly easy task to back track the chain of events and connect the dots that led to the outcome. Here the name of the game is not to describe the uncertainty and complexities surrounding the judgment task but to discount all of the prevailing and latent factors and make dispositional attributions—that is, assign blame—to the individuals involved. Another courtroom ploy that takes advantage of outcome knowledge is the purposeful introduction of extralegal testimony or disallowed information with the full understanding that the jury, even though it is instructed to disregard the information, will be unable to undo its effects and return to their earlier mental state.

Also working to the lawyers’ advantage is the notion of a moral or just imperative. Because of the strong sense of personal responsibility that medical training engenders, it is not uncommon for practitioners to be placed in a position where they are expected to accept more than a fair share of the blame for the sins of omission and commission of everybody else in the greater sociomedical system in which the harmed patient was placed. To assuage the anger and grief of victims and their families in coming to terms with their loss, it seems much more satisfying, but much less enlightening, to be able to point to a responsible individual than to a complex web of system factors which are relatively insignificant in and of themselves but horrific in their combination.16 Medical outcomes that go awry frequently get imbued with moral meaning and notions of just reciprocation. In the aftermath of complex events leading to devastating consequences, we simplify more than we should in seeking understanding and expect to find a level of equivalence with respect to magnitude of consequences and magnitude of causes.17 For example, compared with suffering the consequences of a nosocomial infection, losing a limb in a wrong site surgery generates a heightened degree of moral outrage and a greater likelihood of attributing the loss to gross incompetence as the primary cause when, in fact, a similar degree of system and individual related factors could be involved in both situations.
OUTCOME BIAS IN ASSESSING DECISION QUALITY

While knowledge of an outcome is part of the set of operations used to demonstrate hindsight bias, it is important to make a clear distinction between hindsight bias and outcome bias. Hindsight bias is the tendency for people with outcome knowledge to exaggerate the extent to which they would have predicted the event beforehand, while outcome bias refers to the influence of outcome knowledge upon evaluations of decision quality. Baron and Hershey\(^7\) conducted five experiments on the evaluation of medical and monetary decisions. They found that outcome information consistently influenced evaluations of decision quality, the competence of the decision maker, and the willingness to let the decision makers make decisions for the subject. Although subjects who were asked felt they should not consider outcomes in making the evaluations, they did in fact do so.

Caplan and associates\(^8\) found an inverse relationship between the severity of outcome knowledge (permanent v temporary injury) and physician judgments on the appropriateness of care. They expressed concern about the use of implicit judgments by designated experts when conducting retrospective analysis. Russo and Schoemaker\(^9\) note that many decision makers have difficulty improving their decision making processes because they irrationally judge everything based on outcomes. Their case scenario is instructive. Managers were asked to choose between one of two potential new products to market. It was given that product A had a 50% chance of succeeding while product B had a 60% chance of succeeding. In both cases success meant a one million dollar profit; failure resulted in no profit. The company chose to market product B which failed. Product A was later marketed by a competitor and it succeeded. Did the company make the right decision? Managers were asked to rate the quality of the decision on a scale from 1 (clearly made the wrong decision) to 7 (clearly made the right decision). Their average rating was 4.4, showing a fair degree of confusion as to the correct decision. Why didn't they all respond with a 7 as any group of rational managers should have rather than let the outcome confuse their thinking about the quality of their decision?

Part of the problem may be cultural. In competitive cultures where the stakes are high and personal responsibility runs deep, successful outcomes are highly valued. Whether one is managing new products or making medical decisions, successful outcomes are more highly valued than simply using correct decision processes along the way. It is unfortunate that the strong impact of an outcome can so bias our thinking about the processes that led to that outcome. Given the loose relationship between process and outcome in complex and dynamic settings, and given the weight carried by outcome knowledge, decision makers are better advised to focus on processes. Battlefield commanders, for example, can have excellent battle plans, good intelligence on enemy intentions, and well trained soldiers, and still suffer humiliating defeats. Sometimes good processes can and do lead to unsuccessful outcomes.\(^10\) When this is the case, it is premature to rush in and dramatically revamp processes. Mediocre processes, on the other hand, are sometimes associated with successful outcomes. When this happens, only fools let the success go to their heads.

ROLE OF RECONSTRUCTIVE MEMORY

Is there a way of explaining how outcome knowledge actually leads well meaning people to question so harshly what they feel others should have known from the outset? Research on reconstructive memory provides some clues. The British psychologist, Sir Frederick Bartlett, was one of the first to stress that normal memory was more reconstructive than reproductive.\(^11\) Bartlett was able to show that short stories that are recalled and passed from one person to the next and so on down the line, or that are repeatedly retold from memory by the same person on subsequent occasions, undergo systematic changes. The changes that the story narratives undergo are in the direction of simplification or transformation to fit pre-existing knowledge or beliefs.

Bartlett used the term “schemata” to describe the pre-existing framework onto which new experience is mapped to form a reasonable and coherent whole when later recalled. For example, he found that European students who were asked to recall narratives taken from an unfamiliar culture altered the details in a manner that enabled them to interpret the story more meaningfully in terms of their own culture. As the story was passed from person to person, it had to be repeatedly assimilated in a manner that made sense to each person’s pre-existing framework. With all the idiosyncratic and alien features discarded, what remained was a simplified version of a story that had been rich in detail and nuance, but was now transformed into a more common tale which the subjects of a given culture found more familiar.

Is it possible that “assimilation to the familiar” might be at play in root cause analysis (RCA) techniques for investigating adverse events in healthcare delivery? To be sure, individual characteristics, the nature of the work and workload factors, communication/information transfer, user/equipment interfaces, the physical environment, the social/organizational environment, and managerial dictum cast a broad net and are good places to start. Thorough investigations of adverse events, however, have to do more than focus on the most commonly recognized or familiar contributing factors. Most RCAs are conducted a considerable time after the event occurs by personnel not directly involved in the mishap—two conditions that favor assimilation to the familiar. Although the presence of hindsight bias and assimilation to the familiar are not sufficient reasons for not performing RCAs or any other retrospective analysis, investigators need to attend seriously to these threats if these techniques are to round up anything more than the usual suspects.\(^12,13\)

As noted by Rasmussen,\(^2\) there also are pragmatic and somewhat arbitrary aspects to investigations of human error and system performance. When system performance is below some specified standard or serious mishaps occur, an effort is made to back track the chain of events and circumstances to find the causes. How far back to go or when to apply the stop rule, however, are open questions. Rasmussen notes that the search for causes typically stops when one comes across one or more factors that are familiar and therefore serve as acceptable explanations, and for which there are available corrections or cures. Since there is no well defined starting point to which one is progressively working backwards through the causal chain, how far back one is willing to search is likely to depend on pragmatic considerations such as resources, time constraints, and internal and external political ramifications.

One area where the changeable nature of memory is of particular concern is eye witness testimony as it relates to accident investigation, clinical malpractice, or contested criminal activity. How reliable are the memories of eye witnesses regarding past events that were purportedly observed or experienced? If memory is reconstructive rather than reproductive, as Bartlett tells us, is there a danger that imperfect memories will be filled in or reconstructed to make sense out of the fragments that are remembered? Is it possible that erstwhile testimony will be given regarding events that, indeed, did not take place? Does outcome knowledge influence the reconstructive process? Can even the way questions are framed influence the reconstructive process?

Research on memory and the reliability of eye witness testimony provides a resounding “yes” to the above questions. Loftus\(^23\) described the problem as follows: “Human
Hindsight bias, outcome knowledge and adaptive learning

REDUCING HINDSIGHT BIAS

Reducing hindsight bias is not easy. It is not as simple as issuing a friendly warning to people likely to commit the offence. Fischhoff found that subjects were unable to ignore the stated outcomes, even when instructed to do so. One possibility comes from research by Slovic and Fischhoff who found that hindsight bias could be reduced when people were asked to consider how other possible outcomes might have occurred. By considering how past events may have turned out differently, subjects become less anchored to the outcome which otherwise dominates the reconstruction of the episode and are more open to other possibilities. Arkes and co-workers have shown that the bias can be eliminated by asking hindsight subjects to state all the reasons why potential diagnoses might be correct. A related technique for those who suffer under omniscient superiors or managers who are habitually “smart after the event” is to withhold announcement of the outcome until they have given their own estimates or predictions regarding impending outcomes. When individuals are forced to write down their own views, their own assumptions, their own uncertainties, and their own trade offs, contingencies, and options, they are better able to appreciate the complexity of the decision process which is messy and riddled with unknowns. Otherwise, the outcome appears inevitable and doubts may be cast on the personnel at the sharp end as to why they could not detect the obvious.

AN ADAPTIVE LEARNING CONTEXT

While most of the attention directed at hindsight bias has focused on its undesirable effects, other theoretically inclined investigators have chosen to place it in the broader context of adaptive learning. Hawkins and Hastie observed that, when the task for subjects to respond as if the outcome was unknown is dropped from the hindsight paradigm, one has “a simple learning from outcome feedback tasks”. Here it would be reasonable to expect subjects to update and alter earlier retrospective probabilities as the new information is provided. There would be no need or desire to label the resultant learning as bias. Because learning from outcome feedback task is implicitly embedded in the hindsight paradigm, it is very difficult for researchers to separate pure hindsight effects from adaptive learning. Campbell and Tesler focused on other adaptive aspects of hindsight bias, specifically in relation to personality differences of subjects. Their research suggested that, for some subjects, the value of appearing smart after the fact may be for purposes of maintaining public esteem while, for others, the value may lie in a more orderly control of their environment where the past is reconstructed to be more congruent with a changing present.

The adaptive learning orientations of Hawkins and Hastie and of Hoffrage and Hertwig take us beyond solely focusing on the downside of hindsight. While not denying the ubiquitous nature of hindsight bias in relation to a variety of real world settings, the adaptive learning approach views it as a somewhat troublesome side effect of an otherwise very fast and proficient process for focusing attention on the acquisition of new data when confronted with complex problems, for simplifying and ordering data used in the creation of mental models for inferring causal relations among the evidence attended to, and for continuously updating knowledge structures and revising mental models. Hawkins and Hastie describe hindsight bias as the inability to recapture an earlier, pre-outcome causal model—one that would have been available had not adaptive information processing mechanisms so quickly updated it on the basis of feedback from the outcome. Such a perspective is compatible with accident investigation thinking that urges investigators to focus on reconstructing the mindset of people as they experienced the events unfolding and why their assessments made sense at the time, rather than using hindsight to work out why they did not make sense. Perhaps we have spent too much time at the retrospective end of the human error tunnel and not enough time in active sense making and moving forward.

CONCLUSIONS

Hindsight bias is a pervasive phenomenon that cuts across a variety of judgment domains and needs to be taken seriously, especially by those engaged in retrospective analyses such as morbidity and mortality conferences, error reporting systems, and root cause analysis. Hindsight bias is not simply a matter of lofty academic curiosity, but has very practical real world implications. When a truly independent second opinion is needed, as in
Given knowledge of an outcome, hindsight bias is the exaggerated extent to which individuals indicate they would have predicted the event beforehand. 

Hindsight bias can have an adverse impact on retrospective investigations of events that cause harm. 

The use of case histories for teaching diagnostics is undermined if students only trace backwards someone else’s decision path rather than experience the complexity of the case prospectively with the outcome unknown. 

Outcome bias, a related phenomenon, is the impact of outcome knowledge on evaluations of decision quality, sometimes enabling mediocre processes to be judged as good and good processes as mediocre. 

Outcome knowledge can influence reconstructive memory processes, casting doubt on the reliability of expert and eye witness testimony in legal proceedings involving accidents and medical malpractice. 

Adaptive learning approaches view hindsight bias as a by-product of rapid information processing mechanisms that update knowledge structures. 

Rather than use hindsight to show where people went wrong, investigators should focus prospectively on foresight and reconstruct the conditions experienced by those actively engaged in sense making at the time.

Hindsight bias can have an adverse impact on medical diagnosis, knowledge of the first opinion raises the likelihood of rendering a contaminated second opinion. Important lessons may go unlearned in educational settings if students are allowed simply to back track someone else’s decision landmarks rather than experiencing the confusion, uncertainty, and complexity of the original decision makers.

Outcome bias is a related knowledge of the outcome phenomenon that frequently operates concurrently with hindsight bias. Some of the literature is confusing because of a failure to make a distinction between the two. Hindsight bias is the tendency for people with outcome knowledge to exaggerate the extent to which they would have predicted the event beforehand, while outcome bias refers to the influence of outcome knowledge on evaluations of decision quality. Knowledge of the outcome has enabled mediocre processes to be evaluated as good, and sound processes as poor. 

The pioneering work of Sir Frederick Bartlett on reconstructive memory—where memory gaps are filled in or reconstructed to make sense of the fragments that are remembered—not only implicates the influence of outcome knowledge but casts suspicion on the reliability of eye witness testimony in legal proceedings involving accidents, medical malpractice, and contested criminal activity. 

The reduction of hindsight bias is not easy to achieve. It takes more than raising awareness of its pervasiveness or instructing subjects to disregard the outcome information. Techniques that force individuals to experience the complexity of the decision path, such as offering reasons for alternative diagnoses, have been shown to be helpful. Those who advocate adaptive learning frameworks, while not denying the reality of hindsight bias, approach it in a broader and more positive context. Accordingly, hindsight bias is an unfortunate by-product of an otherwise rapid and proficient process for selecting and processing information as we create or revise our mental models to make sense of the events we experience—a process so well ingrained that it precludes the ability to recover the antecedent, pre-outcome causal state. The adaptive learning framework is compatible with the view that the key to investigations of adverse events is not to focus on where people went wrong but to understand how their decisions and actions made sense at the time.