The problem with ‘5 whys’

Alan J Card

BACKGROUND

The ‘5 whys’ technique is one of the most widely taught approaches to root-cause analysis (RCA) in healthcare. Its use is promoted by the WHO,1 the English National Health Service,2 the Institute for Healthcare Improvement,3 the Joint Commission4 and many other organisations in the field of healthcare quality and safety. Like most such tools, though, its popularity is not the result of any evidence that it is effective.5–8 Instead, it probably owes its place in the curriculum and practice of RCA to a combination of pedigree, simplicity and pedagogy.

In terms of pedigree, ‘5 whys’ traces its roots back to the Toyota Production System (TPS).9 It also plays a key role in Lean10 (a generic version of TPS) as well as Six Sigma,11 another popular quality improvement (QI) methodology. Taiichi Ohno describes ‘5 whys’ as central to the TPS methodology:

The basis of Toyota’s scientific approach is to ask why five times whenever we find a problem ... By repeating why five times, the nature of the problem as well as its solution becomes clear. The solution, or the how-to, is designated as ‘1H.’ Thus, ‘Five whys equal one how’ (5W=1H). (ref. 9, p. 123)

This quote also makes the case for the technique’s simplicity. Asking ‘why’ five times allows users to arrive at a single root cause that might not have been obvious at the outset. It may also inspire a single solution to address that root cause (though it is not clear that the ‘1H’ side of the equation has been adopted as widely).

‘5 WHYS’ AS A TEACHING TOOL

The pedagogical argument for ‘5 whys’ is that it creates an ‘aha moment’ by revealing the hidden influence of a distant cause, which illustrates the importance of digging deeper into a causal pathway. This quick and easy learning experience can be a powerful lesson in systems safety and QI.

Possibly the most famous ‘5 whys’ case study is one that focuses on efforts to preserve the Washington Monument.12 13 Details vary slightly depending on the source, but it usually looks something like this:

Problem: The Washington Monument is deteriorating

1. Why? Harsh chemicals are being used to clean the monument
2. Why? The monument is covered in pigeon droppings
3. Why? Pigeons are attracted by the large number of spiders on the monument
4. Why? Spiders are attracted by the large number of midges at the monument
5. Why? Midges are attracted by the fact that the monument is first to be lit at night.

Solution: Turn on the lights one hour later.

This is a great teaching example because the ‘root cause’ is so unobvious and intuitive. Who would think, before exploring the issue in depth, that lighting choices could endanger a marble monument? But, as is so often the case, reality is messier than this simple illustration.

Joel Gross12 investigated the foundation of this example and discovered that many of the details are incorrect. And, crucially, the broader story it tells is incomplete.

In terms of the story’s details, the monument is an actual event actually the Lincoln Memorial, and it was not being damaged by the use of harsh chemicals. The real culprit was simply water. Pigeons were not an issue at all, and while there were ‘tiny spiders’ (ref. 14, p. 8) at the memorial, they were not a major problem. Instead, most of the cleaning was necessary because swarms of...
The problem with…

midges were dazzled by the lights and flew at high speed into the walls of the memorial, leaving it splattered with bits of the insects and their eggs.\textsuperscript{12, 14}

But that only speaks to the details that were described. The analysis is also incomplete in a number of more important ways. For instance, it only addresses one potential source of deterioration: cleaning water.

The first ‘why’ could just as easily have tackled other causes, such as rain or acid rain (a significant concern at the time), rising damp, erosion from windborne particles or damage from freeze-thaw cycles.\textsuperscript{15}

Or, if the goal had been to prevent harm to future monuments, the first ‘why’ could have focused on the use of marble as a building material, the choice of building site, etc.

However, the most important problem with this example is that, while the solution was ‘effective’ in one sense, it still failed:

Messersmith [the consultant entomologist who worked on this project] thought that because the insects swarmed only at sunset, a one-hour delay in turning on the monument lights would go far in solving the problem. The technique worked, reducing the number of midges in the monuments by about 85 percent.

‘But tourists who had driven hundreds of miles to have their photographs taken at the monuments were not happy,’ he said. ‘They complained every day, and the lights went back on.’\textsuperscript{16}

The logic of the solution was sound, as far as it went. But it was predicated on an incomplete understanding of the broader system, its stakeholders and the purpose of the monument itself. If anything, this window on the complexity of real-world problem solving adds to the value of this teaching example. If the first ‘aha moment’ is followed by this second one, trainees will not only learn that distal causes can have unexpected outcomes, but also that systems thinking requires both depth and breadth of analysis.

THE PROBLEM WITH ‘5 WHYS’ IN RCA

‘5 whys’ has been the subject of a number of caveats and critiques. For instance, Minoura, one of Ohno’s successors at Toyota, highlights the potential for users to rely on off-the-cuff deduction, rather than situated observation when developing answers, as well as difficulty in prioritising causes, if multiple ‘5 whys’ are used.\textsuperscript{17} Mark Graban, a thought leader in the Lean community, points out that ‘5 whys’ is just one component of what should be a far more comprehensive problem-solving process.\textsuperscript{18} And Serrat clarifies that users should not feel constrained by the arbitrary number in the tool’s title: more, or fewer, than five ‘whys’ may be required.\textsuperscript{19}

But the real problem with ‘5 whys’ is not how it is used in RCA, but rather that it so grossly oversimplifies the process of problem exploration that it should not be used at all. It forces users down a single analytical pathway for any given problem,\textsuperscript{13} insists on a single root cause as the target for solutions\textsuperscript{9, 13, 20} and assumes that the most distal link on the causal pathway (the fifth ‘why’) is inherently the most effective and efficient place to intervene.

A single causal pathway

A credible ‘5 whys’ approach to a wrong patient medication error might look like this (adapted from Battles et al.):\textsuperscript{21}

\textbf{Incident: Wrong patient medication error}  
\begin{itemize}
  \item Why? Wristband not checked
  \item Why? Wristband missing
  \item Why? Wristband printer on the unit was broken
  \item Why? Label jam
  \item Why? Poor product design
\end{itemize}

But another team could easily come up with five wholly different and equally valid ‘whys’. And any single string of ‘5 whys’ can provide only a blinkered view of the complex causal pathway that led to the incident. This is illustrated by figure 1, a causal tree diagram (or, more accurately, a ‘causal and contributing factors tree diagram’) depicting the underlying issues that gave rise to the adverse event.

It is clear from the tree diagram that the causal pathway related to the wristband printer is neither the only relevant cause of the incident nor indisputably the most important. A serious effort to solve the myriad problems that gave rise to this incident would have to tackle a number of other causal pathways as well.

These might include pathways related to a maladaptive workplace culture,\textsuperscript{22–25} clinical and information technology (IT) staffing, orientation of agency staff and the absence of a forcing function\textsuperscript{26} to ensure that patients are properly identified before medication is administered. It could also include a focus on improved infection control and better preparedness for infectious disease outbreaks. Solutions based on the ‘5 whys’ in the example above would leave all of these issues unaddressed.

There is also no objective or reliable means of mapping out the causal pathway, which is a critical failing when only one pathway will be examined. Consider the variant below, which follows essentially the same causal reasoning as the first example:

\textbf{Incident: Wrong patient medication error}  
\begin{itemize}
  \item Why? Wristband missing
  \item Why? Wristband printer on the unit was broken
  \item Why? Healthcare system purchased an unreliable printer
  \item Why? Poor process for evaluating and purchasing ‘non-clinical’ equipment
  \item Why? Equipment deemed ‘non-clinical’ is not seen as safety-critical
\end{itemize}

This version skips the step of asking why the wristband was not checked and moves directly to asking...
why it was not there. It also sticks to the high-level issue of the printer being broken, without delving into the details of the label jam. ‘Skipping’ these questions allows the analysis to go deeper because it leaves more ‘whys’ available. This example also maintains a focus on issues within the organisation, rather than the design of the printer. This would lead to very different solutions.

But because this approach skips past the question of why the wristband was not checked, it closes the door to questions about other reasons why it was not checked. In figure 1, this would include the lack of a forcing function. But in another scenario, it might include a desire to avoid waking the patient; an unreadable wristband (eg, smudged, crinkled or occluded); the lack of a label on the medication; confusion caused by multiple wristbands; lack of trust in the wristband data due to frequent errors or any of a number of other causes.

Users could also go down an entirely different causal pathway. An equally reasonable ‘5 whys’ for this incident could look like this:

Incident: Wrong patient medication error
Why? Patients with similar names in the same room
Why? Not feasible to try ‘juggling beds’

**Why?** Not enough nurses to deal with the influx of patients

**Why?** Nurses affected by an outbreak of norovirus

**Why?** Poor adherence to time-consuming infection control interventions

**Why?** A culture of ‘just get the job done’

There are many ‘correct’ ways a team might use ‘5 whys’ to assess even this one incident. And it is unlikely that any two teams would independently arrive at exactly the same results. This subjectivity is critically important because ‘5 whys’ focuses on only one root cause at the end of one causal pathway. More sophisticated practice in the use of ‘5 whys’ might produce two causal pathways, focusing on the main service failures uncovered, rather than the event itself (ie, a set of ‘5 whys’ for ‘wristband not checked’ and another for ‘verbal identification failure’). But this is not how use of the tool has generally been taught in the healthcare industry.

More sophisticated practice in the use of ‘5 whys’ might produce two causal pathways, focusing on the main service failures uncovered, rather than the event itself (ie, a set of ‘5 whys’ for ‘wristband not checked’ and another for ‘verbal identification failure’). But this is not how use of the tool has generally been taught in the healthcare industry. And even this unusually thorough approach would identify only 2 of the 30 causal pathways shown in the tree diagram.

**A single root cause**

Forcing users down a single causal pathway should be disqualifying by itself. But ‘5 whys’ narrows the scope
for improvement even further by insisting that risk control efforts must focus on a single root cause for each causal pathway. In the first healthcare example above, for instance, the root cause would be 'poor product design', and this would serve as the sole target for improvement efforts.

But accidents are seldom the result of a single root cause.33 So focusing exclusively on one (or even a few) arbitrarily determined 'root causes' is not a reliable method for driving improvement—especially in a system as complex as healthcare. As Wieman and Wieman wrote: “Unfortunately … restricting the number of variables [considered] in a complex system only results in an increased potential for errors of omission” (ref. 34, p. 117).

How much might be omitted when using ‘5 whys’? The tree diagram for our example uncovers more than 75 whys (causes and contributing factors), each of which is a potential target for action to reduce the risk of a recurrence. The ‘5 whys’ approach would identify only one (or possibly two) root cause as target for action. At best, this represents <3% of the opportunities for improvement identified using the tree diagram.

**Targeting only the most distal cause**

Not only are users of ‘5 whys’ limited to one root cause per causal pathway, but they are also limited to selecting only the most distal cause (conventionally, the fifth 'why'). There is, however, no logical reason to assume that this is always the most effective or most efficient target for intervention.

Actually, if it were possible to magically place a 100% effective risk control at any one point on the tree diagram, it would be best used on a proximate cause. For instance, making it impossible to administer medication without checking the wristband would render all the more distal causes moot for the purpose of preventing a recurrence. And, while 100% effective risk controls are seldom available, an action plan that includes a proven intervention like a well-designed bar-code reader with a forcing function for patient identification (ID) is more likely to prevent another serious 'wrong patient' medication error than switching to a well-designed printer.

This is not to suggest that more distal causes are not appropriate targets for improvement efforts. In the example presented in *figure 1*, for instance, there is clearly a profound need to change the culture from one that is task-oriented and sometimes hostile to one that is outcomes-oriented and psychologically safe. The pervasive impact of such a culture change would be far more important than merely reducing the risk that this particular incident might recur; it would influence almost every quality and safety issue in the organisation.

And, in contrast to that powerful-but-difficult lever for shifting outcomes, sometimes more distal causes represent ‘low-hanging fruit’ that can be addressed while a more proximate solution is in the works. In *figure 1*, educating patients about why clinicians will be constantly asking them to identify themselves would be far from foolproof. But it *would* be fast, cheap and easy. And it might reduce an important barrier to best practice in verbal identification.

Appropriate targets for intervention may occur anywhere along the causal continuum and on any causal pathway. And efforts to improve safety and quality will often require more than one intervention targeting more than one underlying hazard. It is useful to identify all the key hazards that gave rise to an incident and ensure that each of these is either addressed or intentionally accepted.35 (See, for instance, the Options Evaluation Matrix.)36 But the use of ‘5 whys’ makes this impossible.

**CONSIDERING THE VIRTUES OF ‘5 WHYS’**

What, then, of the virtues of ‘5 whys’? Are the issues above redeemed by the tool’s simplicity and pedigree?

**Simplicity**

Simplicity is a complicated virtue when it comes to the frameworks, tools and techniques of QI. For instance, the conceptual simplicity of the Plan-Do-Study-Act (PDSA) framework is one of its main selling points, but it may also lead organisations to underestimate the messy work involved in applying PDSA to real-world problems.37

But, as this paper has shown, the ‘5 whys’ approach has clearly overshot the mark: it is not simple, but simplistic. It is, as Leveson describes, “… perhaps the most simplistic [accident analysis technique] and … leads to the least amount of learning from events”.13

Charles Vincent famously called for RCA to serve as “a window on the system”.38 If that is the goal, then ‘5 whys’ is doomed to fail. It purposely discards the vast majority of what might be learned about the system being interrogated.

The reality of most healthcare processes and systems is that we face classic design problems: problems that are highly contextualised and often “ill-defined, ill-structured, or ‘wicked’” (ref. 40, p. 224). A ‘5 whys’ analysis ignores this. Most of the causal pathways that led to an event are amputated from the start, and consideration of those that remain is limited to a single root cause.

This creates a toy problem in which it is assumed that simple optimisation of one, or at most a handful of variables will lead to improvement, without any

---

1 A toy problem is: “A deliberately oversimplified case of a challenging problem used to investigate, prototype, or test algorithms for a real problem. Sometimes used pejoratively.” 62

---

need to consider the rest of the system. This flies in the face of everything we know about solving problems in complex adaptive systems like healthcare.

Pedigree

The positive reputation enjoyed by TPS/Lean provides an aura of credibility for ‘5 whys’. But how applicable is this to the question at hand? The reputation of TPS/Lean was built in a very different context. And the use of ‘5 whys’ as an RCA tool is by no means the same thing as the use of the full TPS/Lean methodology.

Healthcare organisations are not automobile factories. And while there is much to be learned from the automotive industry and other high-reliability organisations (HROs), healthcare delivery will never be truly comparable to automobile manufacturing. Despite efforts in the healthcare industry to adopt the tenets of HROs, current practice provides recommended care only about 70% of the time. And the percentage of hospital patients who experience an adverse event may be as high as 25–33%. Reliability in the healthcare industry can improve, and indeed it has (cf. ref. 52). But healthcare is far more complex than automobile manufacturing, and takes place amid processes and systems that are woefully underdesigned in comparison to a modern factory. Further, the safety and quality workforce in healthcare is only beginning to move towards professionalisation and often lacks formal training in engineering, human factors, ergonomics or similar domains.

As a result, approaches developed for solving problems in the automotive manufacturing context may not be as effective in the healthcare arena. And, indeed, the evidence base for the use of Lean/TPS in healthcare is weak and increasingly negative. It is also important to differentiate between the use of ‘5 whys’ as a QI method and the use of Lean/TPS as a QI methodology. Though the two are sometimes conflated in both the literature and practice, they are by no means equivalent. And the use of ‘5 whys’ in healthcare RCA is not typically part of a full-scale Lean management approach.

If the use of ‘5 whys’ does not imply the adoption of Lean, and if the evidence to date does not support the effectiveness of Lean in healthcare in any case, there is little reason to be swayed by the pedigree argument.

Using ‘5 whys’ undermines an already weak RCA process

A recent article by Peerally et al describes a number of important weaknesses in healthcare RCA practice. Some of these have been explored above (eg, focusing on a single root cause or a small handful of them; and poor quality investigations), but it also raises a number of other issues, such as misuse of the RCA process to pursue (or avoid) other agendas; failure to support feedback loops and double-loop learning; a focus on individual and isolated incidents; a confused approach to blame; the ‘problem of many hands’ (see also ref. 63) and the shortfalls of a retrospective approach.

The authors also note that RCA often results in poorly designed and/or poorly implemented risk controls. While the goal of learning from incidents is to reduce risk and prevent future harm, the actual tools and techniques of current practice focus exclusively on diagnosing problems; they provide no direct support for prescribing and managing treatments for the organisational pathologies they uncover.

Some organisations have adopted the PDSA approach to continuous improvement of their risk control actions. But this is a high-level framework, akin to the scientific method. And like the scientific method, it must be implemented through an appropriate set of tools and techniques to produce reliable results. Although a handful of such tools have been introduced in recent years, they have not yet been adopted as the current standard of practice.

Perhaps as a result, there is little evidence to suggest that current practice in RCA improves outcomes.

These challenges do not mean that RCA is never worthwhile; it certainly can be a source of important learning and improvement. But it does mean that we cannot afford to compound these problems through the use of an RCA tool that is so deeply and fundamentally flawed. Other more systems-focused techniques, such as fishbone diagrams, lovebug diagrams, causal tree diagrams, Causal Analysis based on Systems Theory (CAST) or even prospective risk assessment approaches, should be considered instead.

CONCLUSION

When used carefully, ‘5 whys’ may play a powerful role in the classroom. It can illustrate both the need for depth (as a positive example) and the need for breadth (as a negative example) when analysing complex problems.

As a tool for conducting RCAs, however, especially in the area of patient safety, the use of ‘5 whys’ should be abandoned. As the (apocryphal) quote goes: “For every complex problem, there is an answer that is clear, simple and wrong.” When it comes to accident investigation, ‘5 whys’ is that answer.

Competing interests None declared.

Provenance and peer review Commissioned; internally peer reviewed.

The problem with...
The problem with...