Improvement and evaluation

Robert L Wears

Correspondence to
Dr Robert L Wears, Department of Emergency Medicine/CSRU, Univ of Florida/Imperial College London, Jacksonville, FL 32209, USA; wears@ufl.edu

Two related papers[1][2] in this issue of *BMJ Quality & Safety* provide interesting insights into the difficulties of evaluating improvement activities, and also illustrate why improvement is so hard. In a carefully crafted set of controlled, interrupted time series experiments, the authors examined the effectiveness in the operating theatre of two popular improvement interventions: standardised procedures and teamwork training. The primary outcomes in both were process measures: the theatre teams' non-technical skills performance, and the count of 'glitches'—omissions, interruptions or other untoward events that disrupted flow and had potential to affect safety or quality. In both experiments, the investigators took care to ensure the interventions were 'owned' by the frontline workers, and not imposed from without by managers disconnected with the realities of the workplace (although this also means that higher level support important for sustainability may have been lacking).

The papers report insufficient evidence to support improved performance from introducing standard operating procedures, even when those procedures were developed and implemented by the frontline staff themselves.[1] However, they also report a partial success, in that, when accompanied by teamwork training, the combination of standard operating procedures and teamwork significantly improved non-technical skills performance.[2] Curiously, in the combined experiment, technical performance as measured by 'glitches' per hour improved in experimental and control groups. Taken as a whole, the two papers suggest an interaction, or synergism, between the two interventions. Standardisation alone was not effective, but standardisation in conjunction with teamwork training, was (although we cannot be certain whether teamwork alone might have been similarly effective).

These two papers make a valuable contribution to the safety and quality literature by showing that the same intervention (standardisation) can be ineffective in one context (without teamwork training) but effective in another (with teamwork). One wonders how many negative reports of quality interventions were negative only because an important effect modifier was missing from the analysis; or conversely, how many positive reports attributed success to the planned intervention, when it was actually facilitated by an unmeasured interaction variable. There is a significant risk here of drawing the wrong lessons from previous work. This is a possible explanation for the heterogeneity that bedevils the safety and quality literature—a confusing patchwork of claims and counterclaims, reports of interventions that worked or failed, or worked here but not there (sometimes even within the same organisation).[3] Systematic reviews of these reports have not helped much; by dealing with context as a nuisance variable and averaging it out, they tend to cast everything in a dim grey light—across the board, most interventions are neutral or dull average at best, further investigation is required.

These papers fall into a well-established evaluation framework that has become an orthodoxy in healthcare: the technical, rational, deterministic and reductionist approach of positivist 'normal science'. The success of this approach in much of science, and the parallel success in industry of its philosophical cousin, statistical process control, has led healthcare into mistaking the map for the territory. Since positivist science has been such a successful lens through which to view aspects of the world, these aspects have been mistaken for the world and anything that does not fit or cannot be accommodated in a positivist paradigm is tacitly presumed to be unimportant or non-existent.

These methods were largely developed for static, engineered, inanimate systems; the paradigmatic model for statistical process control is the assembly line. They are approaches suitable to machines—there are seldom interactions among

Linked

CrossMark

The Health Foundation Inspiring Improvement

BMJ

components, it is possible to change only one thing at a time, as a change in one part does not produce a consequent change in another.

However, healthcare systems are not assembly lines. They are complex, intractable, sociotechnical systems and[4–6] organic rather than engineered. Their basic 'physics' is poorly understood at best. They do not simply accept change (eg, interventions), but adapt and reconfigure themselves in response to it; those adaptations reverberate and ramify throughout the system via positive and negative feedback loops with varying delays. These interactions among components are more important than the components themselves; the behaviour of one component depends in part on the behaviour of others, and the evolving cycles of reciprocal action and reaction reshape the universe of possibilities.[7 8] This makes systems path dependent; the past trajectory of changes, reactions, and interactions influences future paths, opening some while closing others.[9] Furthermore, sociotechnical systems are composed at least in part of sentient beings, so how those actors in the system understand and interpret interventions in context, and develop strategies to manage or integrate them within existing workflow, have strong influences.

These properties make it impossible to change only one thing,[10 11] and difficult to predict the overall effect of changes by 'summing' across the individual effects.[7] Thus, interventions in a complex sociotechnical system produce a chain of consequences that extend over time and cannot be fully anticipated. Such systems cannot be directly controlled in the Taylorist, rationalist way that managers or regulators would like; and evaluations of interventions in such systems can never be 'one and done', but must always be formative rather than summative.

The problem is exacerbated when the intervention itself is a complex social one.[12] In the two papers discussed here, teamwork training is clearly a complex social intervention, but what about standard operating procedures? Standardisation is often viewed as a purely objective, technical exercise, but this is a misconception.[13] However, objective, rationalised, complete and internally consistent a set of standardised procedures might be, their development, interpretation and application are social processes, subject to the context, history, politics and goals of actors in the system.[14] In addition, there are inevitably gaps between the imagined world of the procedures and the real world of work,[15] and conflicts among competing goals; both must be recognised, negotiated and resolved in action by workers in a community of practice. Finally, the cycle of adaptations set in motion by the intervention can feed back onto the original intervention itself, so that it also changes with time, triggering yet another cycle of adaptations.

Although complex sociotechnical systems cannot be directly controlled, all is not lost, because they can be influenced.[8] Interventions may not lead directly to the desired behaviours, but they can 'set the stage' to enhance and sustain the emergence of those behaviours.[16] This realisation will require us to modify our approach to both improvement and its evaluation. It will require accepting a broader range of sciences and methodologies as admissible; abandoning many of the Taylorist principles that have informed improvement efforts;[17] and fundamentally re-examining the Newtonian-Cartesian assumptions that underlie them.[18]

Similarly, we will have to expand our evaluation methods to move beyond a certain methodological fetishism[19] aimed at answering the 'horse race' question "Does A work better than B?" and adopt more nuanced methods[20–22] aimed at a more complex set of questions: "Which works, how, why, for whom, to what extent and in what context?" These questions are often best addressed by qualitative, ethnographic methods aimed at providing a 'thick description' in a case study of an improvement effort.[23–25] The value of this type of approach has been shown by careful, theory-driven studies of how and why initiatives are successful[26]: for example, discovering that the theory of improvement motivating a project at its beginning was not the way in which improvement actually, eventually occurred; or illuminating tensions and paradoxes in contrasting understandings of interventions.[27]

However, progress in this area is haunted by a difficult question: why is it that safety and quality in healthcare has been so strongly wedded to rationalist, Taylorist, Cartesian-Newtonian thinking about the nature of clinical practice, and how to improve it? Three factors supporting this marriage may be difficult to overcome. First, it offers the comforting modernist illusion that the muscular application of science can at last tame risk, uncertainty, and disorder, leading to a better, safer, more controllable world.[28] Second, it offers a satisfying explanation for drawing meaning out of the inevitable failures that still must occur,[29] while simultaneously not threatening those in power.[30] And finally, it supports a long-standing secular trend increasing the power and influence of a technocratic elite[18] of scientific-bureaucratic managers[31] that accompanies the progressive industrialisation of healthcare.[32 33] Ironically, the external pressures on healthcare to achieve the precision, safety and efficiencies of linear production systems is driving some very counter-productive behaviours and undermining our desired goals.

REFERENCES

1 Morgan L, New S, Robertson E, *et al.* Effectiveness of facilitated introduction of a standard operating procedure into routine processes in the operating theatre: a controlled interrupted time series. *BMJ Qual Saf* 2015;24:120–27.

2 Morgan L, Pickering SP, Hadi M, *et al*. A combined teamwork training and work standardisation intervention in operating theatres: controlled interrupted time series study. *BMJ Qual Saf* 2015;24:111–19.

3 Davidoff F. Heterogeneity is not always noise. *JAMA* 2009;302:2580–6.

4 Waterson P. Sociotechnical design of work systems. In: Wilson JR, Corlett N, eds. *Evaluation of human work*. 3rd edn. London, UK: Taylor & Francis, 2005:769–92.

5 Eason K. Afterword: the past, present and future of sociotechnical systems theory. *Appl Ergon* 2014;45:213–20.

6 Kleiner BM. Macroegonomics: work system analysis and design. *Hum Factors* 2008;50:461–7.

7 Jervis R. *System effects: complexity in political and social life*. Princeton, NJ: Princeton University Press, 1998:328.

8 Axelrod R, Cohen MD. *Harnessing complexity: organizational implications of a scientific frontier*. New York, NY: Basic Books, 2000:184.

9 Vaughan D. System effects: on slippery slopes, repeating negative patterns, and learning from mistake? In: Starbuck HW, Farjoun M, eds. *Organization at the limits: NASA and the Columbia Accident*. London, UK: Blackwell, 2005:41–59.

10 Thomas L. On meddling. *N Engl J Med* 1976;294:599–600.

11 Sterman JD. System dynamics modeling: tools for learning in a complex world. *California Manag Rev* 2001;43:8–25.

12 Davidoff F. Improvement interventions are social treatments, not pills. *Ann Intern Med* 2014;161:526–7.

13 Wears RL. Standardisation and its discontents. *Cogn Technol Work* 2014. doi:10.1007/s10111-014-0299-6 [epub ahead of print 26 Sep 2014].

14 Høyland S, Aase K, Hollund JG, *et al*. What is it about checklists? Exploring safe work practices in surgical teams. In: Bieder C, Bourier M, eds. *Trapping safety into rules: how desireable or avoidable is proceduralization*. Farnham UK: Ashgate, 2013:121–38.

15 Hollnagel E. Why is work-as-imagined different from work-as-done? In: Wears RL, Hollnagel E, Braithwaite J, eds. *Resilience in everyday clinical work*. Farnham, UK: Ashgate, 2015 (in press):249–64.

16 Hilligoss B. Selling patients and other metaphors: A discourse analysis of the interpretive frames that shape emergency department admission handoffs. *Soc Sci Med* 2014;102:119–28.

17 Berwick DM. Improvement, trust, and the healthcare workforce. *Qual Saf Health Care* 2003;12:448–52.

18 Wears RL, Hunte GS. Seeing patient safety 'Like a State'. *Saf Sci* 2014;67:50–7.

19 Greenhalgh T, Howick J, Maskrey N. Evidence based medicine: a movement in crisis? *BMJ* 2014;348:g3725.

20 Berwick DM. The science of improvement. *JAMA* 2008;299:1182–4.

21 Pawson R, Tilley N. *Realistic evaluation*. London, UK: Sage Publications, Ltd, 1997:235.

22 Greenhalgh T, Russell J. Why do evaluations of eHealth Programs fail? An alternative set of guiding principles. *PLoS Med* 2010;7:e1000360.

23 Flyvbjerg B. Case study. In: Denzin NK, Lincoln YS, eds. *Sage handbook of qualitative research*. 4th edn. Thousand Oaks, CA: Sage, 2011:301–16.

24 Geertz C. Thick description: toward an interpretive theory of culture. In: *The interpretation of cultures: selected essays*. New York, NY: Basic Books, 1973:3–30.

25 Leslie M, Paradis E, Gropper MA, *et al*. Applying ethnography to the study of context in healthcare quality and safety. *BMJ Qual Saf* 2014;23:99–105.

26 Dixon-Woods M, Bosk CL, Aveling EL, *et al*. Explaining Michigan: developing an ex post theory of a quality improvement program. *Milbank Q* 2011;89:167–205.

27 Greenhalgh T, Potts HW, Wong G, *et al*. Tensions and paradoxes in electronic patient record research: a systematic literature review using the meta-narrative method. *Milbank Q* 2009;87:729–88.

28 Dekker SWA, Nyce J, Myers D. The little engine who could not: "rehabilitating" the individual in safety research. *Cogn Technol Work* 2013;15:277–82.

29 Dekker SWA. The psychology of accident investigation: epistemological, preventive, moral and existential meaning-making. *Theor Issues Ergon Sci* 2014. doi:10.1080/1463922X.2014.955554 [epub ahead of print 14 Oct 2014].

30 Dekker SWA, Nyce JM. There is safety in power, or power in safety. *Saf Sci* 2014;67:44–9.

31 Harrison S, Moran M, Wood B. Policy emergence and policy convergence: the case of 'scientific-bureaucratic medicine' in the United States and United Kingdom. *Br J Politics Int Relations* 2002;4:1–24.

32 Starr P. *The social transformation of American Medicine: the rise of a sovereign profession and the making of a vast industry*. New York, NY: Basic Books, 1983:528.

33 Kleinke JD. The industrialization of health care. *JAMA* 1997;278:1456–7.