BMJ Quality & Safety Online First, published on 18 June 2015 as 10.1136/bmjqs-2015-004431

The problem with checklists

Ken Catchpole,[1] Stephanie Russ[2]

[1]Department of Surgery, Cedars-Sinai Medical Center, Los Angeles, California, USA
[2]University of Aberdeen, Aberdeen, UK

Correspondence to
Dr Ken Catchpole, Department of Surgery, Cedars-Sinai Medical Center, Los Angeles, CA 90048, USA; Ken.Catchpole@cshs.org

Accepted 27 May 2015

'The Problem with…' series covers controversial topics related to efforts to improve healthcare quality, including widely recommended but deceptively difficult strategies for improvement and pervasive problems that seem to resist solution.

Since the seminal studies by Gawande and colleagues[1] and Pronovost *et al*,[2] checklists have become the go-to solution for a vast range of patient safety and quality issues in healthcare. Some see them as a quick and obvious solution to a relatively straightforward problem. For others, they illustrate a failure to understand and address the complex challenges in patient safety and quality improvement. Indeed, successes[3] and failures[4–6] illustrate an underlying difficulty with understanding precisely why checklists work in some cases but not in others. A recent viewpoint summarises the varying applications of checklists in aviation and healthcare, reflecting upon the dangers of making assumptions about their 'ubiquitous utility'.[7] This provided a timely "The Problem with…"[8] opportunity, in which we consider the narratives that often surround the complex challenges faced in designing and implementing a successful checklist, and the science used to explore it.

'A SIMPLE IDEA FROM OTHER INDUSTRIES…'

The apparent simplicity of a checklist is understandingly tempting, with some narratives suggesting that their adoption can be used to effectively address what would appear to be intractable, complex and potentially painful systems issues. However, this simple narrative does not always reflect an understanding of the problems needing to be solved, how best to solve them or indeed the intricacies surrounding the implementation, use and impact of such a simple looking tool. More likely, what we face in introducing a checklist is a rather more complex story of gains and losses, procedural interactions and sociocultural balances (see table 1). This 'simple' versus 'complex' narrative can also be seen in the frequent aviation analogies, which imply that checklists prevented accidents ('simple'), while omitting to mention the critical design changes that were also necessary ('complex'). For example, on the B-17 aircraft, flap and gear levers required redesign as they were easily confused, critically positioned and thus predisposing to accidents.[10]

We have often compared healthcare checklists and their evidence base with checklists used in other industries, but there are some important qualitative differences. For example, as the companion viewpoint also observes, healthcare checklists do not always share design features with their aviation counterparts. For an Airbus A319 (figure 1), a single laminated gatefold (four sides of normal A4 paper) contains the 13 checklists for normal and emergency operations. Tasks range from 2 (for cabin fire checklist) to 17 (for before take-off checklist), with an average of seven per checklist. Each task is described in no more than three words and can be checked immediately, with usually a single word of confirmation. It has no check boxes, does not require signature and is designed to be used by one person, with specific checklists performed aloud. In contrast, the Centers for Disease Control and Prevention central line-associated blood stream infections checklist[11] has 18 tasks, with no less than 4 word descriptors (and up to 22 words), and describes non-procedural tasks that need to be completed over several minutes (and hours), which cannot be 'checked' (eg, 'empower staff'). The WHO safer surgery checklist (first edition)[12] has 21 tasks (7+7+7), with wording ranging from 2 to 16 per task, and involves several people simultaneously. Some tasks are easily checked and completed, while some require discussion and some cannot be 'checked'.

One feature of checklists in healthcare, in comparison to most other industrial uses, is that they increasingly

To cite: Catchpole K, Russ S. *BMJ Qual Saf* Published Online First: [*please include* Day Month Year] doi:10.1136/bmjqs-2015-004431

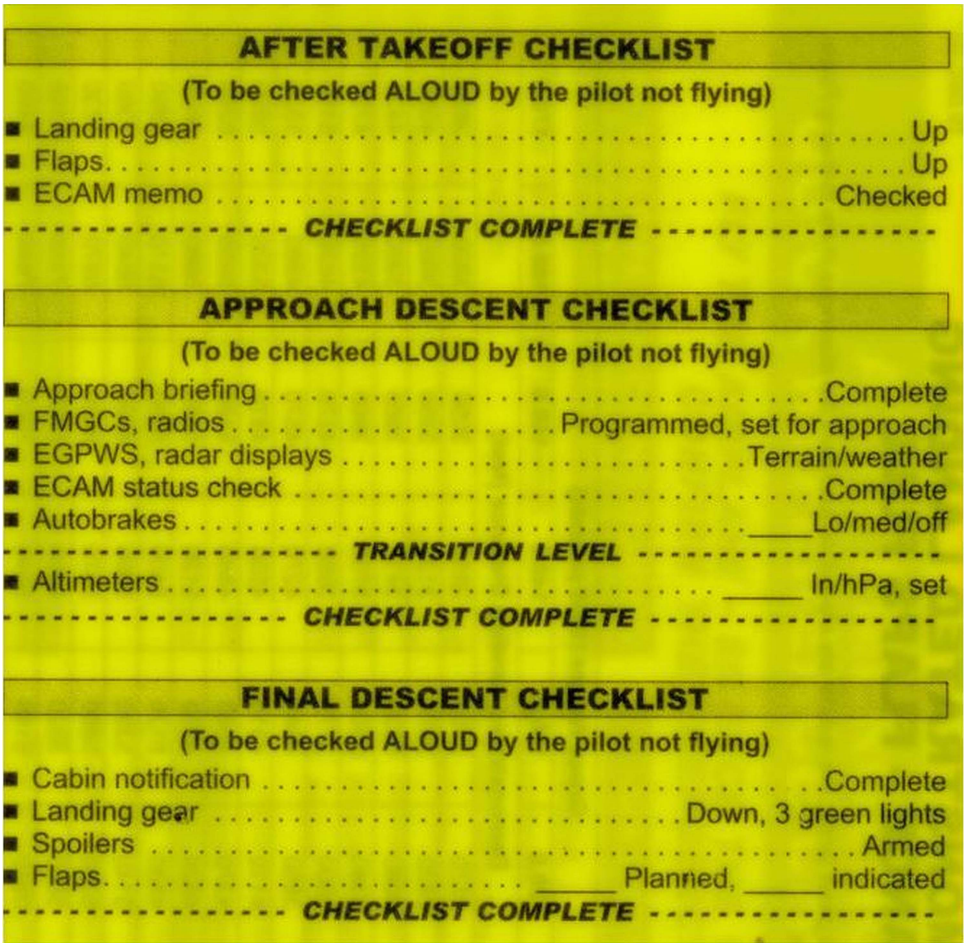
feature items intended to promote communication and teamwork (eg, introductions, discussion of patient risk factors, concerns and so on), in addition to straightforward categorical checks (eg, have hands been washed, has consent been obtained and others). If used in the right way, they can indeed assist in the change of communication patterns and specific coordination tasks (such as 'call outs', task allocation and task visibility).[10] However, creating an opportunity for a more general team talk is not a traditional feature or necessarily a particular strength of checklists. In fact, authentic checklist completion will rely on good communication and teamwork in the first place, which is not always the case. This use as a 'teamwork and communication tool' therefore goes beyond the classic uses of an 'aviation style' checklist.

Team-related checklists have further difficulties because there is often no single calm moment where everyone can be involved, and not every user will experience a direct benefit despite the effort and initiative required for their use. Furthermore, the more complex sociological and cultural challenges, such as power distance, hierarchy and perceptions of professionalism, will often continue to dominate interactions regardless of the design and implementation of the checks. Introducing a checklist reliant on communication without consideration of these processes undermines the sociocultural underpinnings of the intervention and the complexity of coordinating a complex team around a single task.

Thus, prominent uses of checklists in healthcare, such as the surgical safety checklist and the central

Table 1 Simple versus complex narratives in checklist rationale, design and use

	Simple narrative	Complex narrative
Rationale	Checklists reduced aircraft accidents	Aircraft accidents were reduced through a range of physical and procedural design changes, of which checklists were one component
	They 'plug the holes' in the Swisscheese[9]	There are many solutions to a problem; checklists may solve problems and/or introduce new ones
	They encourage safer behaviour	They may encourage 'mindless' checking, promote automaticity and discourage conceptual thinking about a task
	They reduce undesirable human variability	Variable human responses address contextual variability, essential for safety system function
	They are 'evidence based'	They are one part of the underlying mechanism of effect, which is often poorly defined
	They are an exemplar of 'systems thinking'	They are used to modify behaviour instead of applying broader systems thinking
	They encourage teamwork and communication	They can help to promote shared awareness and team discussions where sufficient team skills and a supportive working environment already exist, but cannot achieve this alone
	They are for checking something is correct	They are for checking something has been done
	They encourage 'pause for thought' or discussion	They are most effective when requiring immediate stimulus-response behaviour
	They reduce the effects of interruptions	They are most error-prone when interrupted
Design	They are a simple piece of paper	They are a complex socio-technical intervention
	They can be developed easily	They take considerable effort to be effective, with many design dimensions to consider
	They are a 'stand-alone' solution	A checklist is part of a wider engineered process, including other checklists
	They define how a task should be performed	The user should be skilled and well practiced at the task. A checklist should assist them in doing it
	They are a simple set of statements and boxes	There are a wide range of design parameters
	They must have a tick box	A tick box is not always necessary and does not guarantee full or proper use
	Text is descriptive of desired performance	The text should be a reminder for a motivated user, already skilled and experienced in the task
	They should be ordered in terms of function (ie, All similar process items together)	They should be ordered in terms of geographical and temporal proximity (ie, tasks done in the same time and space)
	They can be used for general tasks ('empower staff')	They should be used for specific tasks ('manual start switch…off')
Use	They can be implemented easily	Implementation is a complex and challenging process that also requires ongoing maintenance
	They are a cost-effective solution	The resources required to implement, perform and maintain a checklist are rarely calculated
	They need to be followed by everyone	A challenging dichotomy arises where experts, who may perceive them as wasteful and patronising, use them as a reminder only. This may or may not be appropriate
	They should always be complied with	A poor design, implementation or context might make it impossible to be compliant
	They should be signed	Signing does not guarantee appropriate use, accountability, compliance or audit accuracy, and can promote 'gaming' and false views of safety
	Lack of professionalism and a culture of safety are the causes of non-compliance	Non-compliance may be because of inappropriate designs, use cases, implementation, training, perceived utility, threats to professional identity/autonomy/expertise, power-play by managers and time/cost burdens incongruent with other system demands
	Their use can be easily and reliably audited	Real-time observations of checklist use typically reveal lower levels of compliance than those suggested by organisational audits. True compliance refers to 'how' checklists are being used, not just 'if', and this should be measured to maximise understanding of the barriers and facilitators to uptake and buy-in



AFTER TAKEOFF CHECKLIST

(To be checked ALOUD by the pilot not flying)

- Landing gear ..... Up
- Flaps ..... Up
- ECAM memo ..... Checked

CHECKLIST COMPLETE

APPROACH DESCENT CHECKLIST

(To be checked ALOUD by the pilot not flying)

- Approach briefing ..... Complete
- FMGCs, radios ..... Programmed, set for approach
- EGPWS, radar displays ..... Terrain/weather
- ECAM status check ..... Complete
- Autobrakes ..... ____ Lo/med/off

TRANSITION LEVEL

- Altimeters ..... _____ In/hPa, set

CHECKLIST COMPLETE

FINAL DESCENT CHECKLIST

(To be checked ALOUD by the pilot not flying)

- Cabin notification ..... Complete
- Landing gear ..... Down, 3 green lights
- Spoilers ..... Armed
- Flaps ..... _____ Planned, _____ indicated

CHECKLIST COMPLETE

Figure 1 Detail of 'Normal' Checklists for Airbus A319 (United Airlines, 19 December 2003).

line bundle, are far from 'simple' and are not that similar to applications of checklists in the industries that inspired their use.

'...THAT LEADS TO BETTER OUTCOMES'

Successful replication of interventions requires us to know why and how they work, that is, to understand the mechanism of effect.[13] A checklist will support a well-defined procedural task that may be vulnerable to problems with memory, task omissions or order. Success may be due to the specific items on the checklist, the non-specific team talking that is promoted, or some other artefacts of the intervention. It is not always clear which is the most important, or how this varies according to context. If it is the 'team talk' that is important, then the items on the checklist and, indeed, the checklist itself, may be inconsequential.[14] Alternatively, if the successful use of a checklist requires the systematic completion of every item, then efforts to ensure compliance should be directed towards the appropriate checking of each item. However, a checklist reliant on teamwork for success may fail despite all the items being followed, because those team skills were insufficient. This can lead to compliance and audit measures that may unfairly penalise (not 'checking all the boxes') while discouraging appropriate use (having a team discussion). Thus the 'evidence' for 'checklists' can quickly become a semantic or mechanistic misattribution and a correlation/causation fallacy.

This confusion with mechanism and assurance is superimposed on the existing dichotomy of the checklist as a 'rule'—which requires everyone to use the checklist, methodically, in the same way, or as an 'aid'—where it may be used differently by experts who are already reliable and proficient. Without direct observation of checklist use, therefore, it is at least open to significant challenge that observed outcomes are related to checklist use or, alternatively, that full compliance may not be necessary to achieve the desired effect. Future studies must seek to address the level and type of compliance (which is unlikely to reside at 100%) required to achieve the observed outcome effects, and understand and define the mechanisms of effect.

The failure to replicate results in healthcare[4 5] demonstrates that at the very least, where success has been observed, it was not only a checklist that created the effect.[15] Much of the variation in reported impact has been attributed to the style of implementation

employed. While we are coming to acknowledge that simply presenting a checklist to users and enforcing strict progression through the items upon it may not necessarily yield performance or outcome benefits, optimal checklist implementation approaches are becoming notoriously difficult to define in healthcare. Favourable implementation conditions are well published (eg, education, local modification, feedback, senior support),[16] however, still do not reliably lead to positive checklist effects.[5] This makes it difficult to draw lessons from the international evidence base that will work at a local level.

Just as checklists were necessary, but not sufficient, in aviation, it seems probable that viewing checklists as an adequate intervention to address the complexities of modern-day medicine in isolation is a somewhat naive viewpoint. More likely, their integration into broader multifaceted team-training 'packages' that incorporate additional team-based interventions (eg, briefings and de-briefings), include a focus on culture and support the skills required to complete the process will likely create a milieu in which a well-designed checklist can work.[14]

LOST IN TRANSLATION?

There is no question that the right checklist, in the right place, with the right design and implementation, can be used enthusiastically by the right people with the right skills and can be highly effective. Yet, in translating checklists from aviation and other industries to healthcare, we may have misunderstood their strengths, failed to design them based on well-established principles and failed to engineer them as a component of a wider socio-technical system. We have made assumptions about their use, effectiveness and 'evidence base' that are readily and easily challenged, and have defined compliance criteria and penalties based on assumptions that may not reflect how they contribute to better outcomes. The superimposition of teamwork and communication—without specifically providing training for those skills, or indeed the sociocultural support for them—further contributes to the difficulties in successful implementation.

A checklist is a complex socio-technical intervention[17 18] that requires careful attention to design, implementation and basic skills required for the task.[19] Understanding and specifying these mechanisms of effect with greater precision would enable us to move beyond the moot 'checklists do/don't work' commentaries. Much as we would like it to be true, the story that checklists are a 'simple, evidence-based solution successfully used in other industries' has many of the characteristics of 'Cargo Cult' or pseudoscience.[20] There is indeed a science to checklists. But, unless we pay attention to the more complex narrative for how they emerged in other industries, including the other changes (to culture, teamwork and design) that accompanied them, we stand little chance of appreciating that science or realising similar benefits in healthcare.

Competing interests None declared.

Provenance and peer review Not commissioned; internally peer reviewed.

REFERENCES

1 Haynes AB, Weiser TG, Berry WR, *et al*. A surgical safety checklist to reduce morbidity and mortality in a global population. *N Engl J Med* 2009;360:491–9.
2 Pronovost P, Needham D, Berenholtz S, *et al*. An intervention to decrease catheter-related bloodstream infections in the ICU. *N Engl J Med* 2006;355:2725–32.
3 de Vries EN, Hollmann MW, Smorenburg SM, *et al*. Development and validation of the SURgical PAtient Safety System (SURPASS) checklist. *Qual Saf Health Care* 2009;18:121–6.
4 Urbach DR, Govindarajan A, Saskin R, *et al*. Introduction of surgical safety checklists in Ontario, Canada. *N Engl J Med* 2014;370:1029–38.
5 Reames BN, Krell RW, Campbell DA Jr, *et al*. A checklist-based intervention to improve surgical outcomes in Michigan: evaluation of the keystone surgery program. *JAMA Surg* 2015;150:208–15.
6 Bion J, Comm W, Richardson A, *et al*. 'Matching Michigan': a 2-year stepped interventional programme to minimise central venous catheter-blood stream infections in intensive care units in England. *BMJ Qual Saf* 2013;22:110–23.
7 Clay-Williams R, Colligan L. Back to basics: checklists in aviation and healthcare. *BMJ Qual Saf* 2015.
8 Shojania KG, Catchpole K. 'The problem with…': a new series on problematic improvements and problematic problems in healthcare quality and patient safety. *BMJ Qual Saf* 2015;24:246–9.
9 Reason J. Human error: models and management. *BMJ* 2000;320:786–70.
10 Roscoe S. *The adolescence of engineering psychology*. The Human Factors and Ergonomics Society, 1997. https://http://www.hfes.org/Web/PubPages/adolescencehtml.html (accessed 19 May 2015).
11 Prevention CfDCa. Checklist for the Prevention of Central Line Associated Blood Stream Infections. 2011. http://www.cdc.gov/HAI/pdfs/bsi/checklist-for-CLABSI.pdf (accessed 19 May 2015).
12 World Health Organisation Surgical Safety Checklist. First Edition. http://www.who.int/patientsafety/safesurgery/tools_resources/SSSL_Checklist_finalJun08.pdf (accessed 19 May 2015).
13 Chopra V, Shojania KG. Recipes for checklists and bundles: one part active ingredient, two parts measurement. *BMJ Qual Saf* 2013;22:93–6.
14 Neily J, Mills PD, Young-Xu Y, *et al*. Association between implementation of a medical team training program and surgical mortality. *JAMA* 2010;304:1693–700.
15 Degani A, Wiener EL. The Human Factors of Flight Deck Checklists: The Normal Checklist. 1990. NASA Contractor Report 177549. National Aeronautics and Space Administration. http://ntrs.nasa.gov/archive/nasa/casi.ntrs.nasa.gov/19910017830.pdf

16 Russ S, Sevdalis N, Moorthy K, *et al*. A qualitative evaluation of the barriers and facilitators toward implementation of the WHO surgical safety checklist across hospitals in England: lessons from the Surgical Checklist Implementation Project. *Ann Surg* 2015;261:81–91.
17 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.
18 Bosk CL, Dixon-Woods M, Goeschel CA, *et al*. Reality check for checklists. *Lancet* 2009;374:444–5.
19 Drury CG. Human factors and ergonomics audits. In: *Handboook of human factors and ergonomics*. 3rd edn. Hoboken, NJ: John Wiley & Sons, 2006:1106–32.
20 Feynman R. Cargo Cult Science. *Surely You're Joking, Mr. Feynman!* New York, NY: WW Norton & Company, 1997:338–46.

BMJ Quality & Safety

The problem with checklists

Ken Catchpole and Stephanie Russ

BMJ Qual Saf published online June 18, 2015

Notes