

# Tickbox

REVIEWED BY KARL G. RULING

IN “THE SAFETY DANCE” in the Summer 2022 *Protocol*, Michael Matthews and Alan Rowe complained about the “checklist safety” mentality—people grasping a particular thing as an indicator of safety without understanding the interplay of things that have to work together to make an event safe. That reminded me of *Tickbox*, a book by David Boyle about pretty much the same problem but by a different name that had received good reviews in *The Guardian*. I bought it and read it. I am now happy to recommend it to *Protocol* readers. It’s an interesting and thought-provoking book—and it may help readers avoid “. . . but I ticked the box!” disasters.

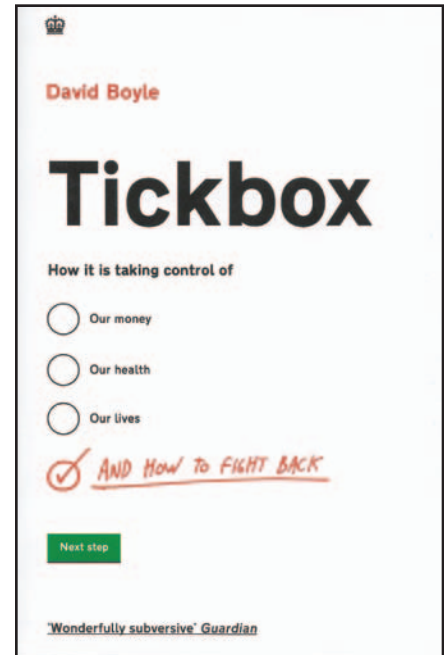
**Tickbox culture attempts to reduce complex human situations down to simple metrics . . . simple steps that can be implemented . . . by any minimally trained, low-paid worker.**

*Tickbox* and “The Safety Dance” hit on the same problem: people simplifying a complex issue, reducing the management of it to a set of little boxes to tick off, thus resulting in a poorly managed or unsafe situation. *Tickbox*, at 262 pages, deals with the issue in much more depth, but both the book and the article make a distinction between lists that help people make sure they haven’t missed something and lists that are designed to obviate thinking. Boyle calls the first type of list a “checklist,” the second a “tickbox.” Checklists are designed

to help people not miss something important; they empower people. Tickboxes look the same, but they are designed to facilitate bureaucracies by replacing human observation and judgment with algorithms. Tickboxes are designed to disempower—to tell people to shut up and tick the boxes. When they are all ticked, all is good! To this “tick the boxes,” Boyle adds the five-point customer surveys we all seem to get after we use any service, and labels this all “tickbox culture.” Tickbox culture attempts to reduce complex human situations down to simple metrics that can be used in management reports, advertisements, and sales meetings—and to boil every

procedure down to simple steps that can be implemented (perhaps not well) by any minimally trained, low-paid worker.

Boyle gives a fascinating history of the development of tickbox and some of the famous names in its development and application. He starts with Frederick Winslow Taylor, who did time-motion studies of workers on the factory floor to develop scientific ways of making assembly work more efficient. He broke down every job into its components, figured out how long each step should take (a simple metric:



time), and then organized pay scales so that those who met the criteria for the most efficient people got the most pay. This system made him famous at Bethlehem Steel, where his improvements by 1901 had the workforce handling three times as much material and earning 60% higher wages than before. Taylor also got himself fired and never worked as an employee again, becoming one of the first professional management consultants.

The five-point scoring system, the Likert Scale, has a history intertwined with mass incarceration and the fire-bombing of Dresden and Tokyo. The scale was developed by the sociologist Renis Likert, who used it to judge the attitudes of Americans about “enemy aliens” during World War II. On the Pacific coast, surveys done by other researchers found people nervous about those of Japanese origin, so Japanese Americans were forced out of their homes into internment camps. On the Atlantic coast, Likert ran his own more nuanced survey using the scale, but also allowing freely written comments. He found people were nervous about

“enemy aliens,” but all people really wanted was the government to keep an eye on anyone suspicious. People of German and Italian descent were not rounded up and imprisoned on the East coast. “To the end of his life, Likert felt that—had he exerted himself more to interpret the Pacific coast results—the Japanese Americans too would have been left alone,” writes Boyle.

Between 13 – 15 February 1945, RAF and USAAF bombers dropped more than 3,900 tons of bombs on Dresden. The resulting firestorm destroyed more than 1,600 acres of the city center and killed an estimated 22,700 to 25,000 people. The following month, in one night, American B-29s rained incendiary bombs on the residential sections of Tokyo, touching off a firestorm that burned a quarter of the city to the ground, killed about 100,000 people, and left a million homeless. These were fabulously expensive air raids. Were they worth it?

To find out, President Franklin Roosevelt set up the Strategic Bombing Survey, with a team of over 1,200 people, including Likert and John Kenneth Galbraith, to interview Japanese and German civilians (how researchers got access to the populations is not clear) to find out how the bombing affected them, with the idea that destroying morale would justify the expense and carnage. Likert’s five-point scale was used but not as the sole tool; the surveys were complicated by their seeming to be no direct Japanese equivalent to “morale.” The director of the project, Paul Nitze, became convinced that the Japanese were close to surrender, the bombing could stop, and events would take their course. Hiroshima and Nagasaki were bombed regardless.

Boyle writes that Robert McNamara “was certainly involved in the Strategic Bombing Survey.” He would become famous later for his tickbox method of assessing success in the Vietnam War: body count. Success was a simple metric: how many dead bodies were left in the field after a battle. Complications were that the Viet Cong tried hard to reclaim their dead and American soldiers got themselves killed while counting the

bodies. Furthermore, “You might just gain yourself more enemies the more you kill,” writes Boyle. That said, a precise number of dead bodies seems to say something, even if it has little real correlation to success in war.

Boyle takes aim at “Key Performance Indicators,” the KPI targets bosses set to prod workers. According to “Goodhart’s Law” (named after the economist Charles Goodhart), any measurement that becomes a target ceases to become a useful measurement. If you tell employees they must make a target or they’ll be sacked, they will do it, but the results might not be what you want. He holds up as an example Wells Fargo Bank, which set targets for employees for cross-selling their products. Bonuses were offered to employees who sold to existing customers, and employees who failed to do so were told that their continued employment would be reviewed. As a result, employees opened false accounts to meet the sales goals. After the news broke in 2016, Wells Fargo was fined \$185 million and 5,300 employees were sacked.

Boyle doesn’t mention the VW dieselgate scandal, but it fits the pattern. Engineers were tasked with designing an engine system that delivered good fuel economy, spunky performance, and that would pass emissions tests. Three tickboxes, and the engineers could tick them all—but not at the same time. The VW TDI engine controller had different mappings for when the car was being driven or tested. The NO<sub>x</sub> emissions during real-world driving were 40 times higher than during testing. The fines and litigation have been immensely expensive.

David Boyle’s most appalling stories of management by tickbox come from the UK’s immigration service. The Windrush scandal was predictable given the decision system put in place. When evidence presented by asylum seekers didn’t match exactly what the tickbox seemed to need, the application went to the bottom of the pile. The algorithm had no place for human intervention, for anything that would take a person’s time to investigate and to consider. Staff working on asylum applications were

incentivized to mark puzzling cases as “non-straightforward,” which put the applicants who had lived in the UK for decades and paid taxes into legal limbo—or into jail.

Boyle’s book is not against all systematic decision-making and management, but you have to be careful about defining what you are managing and how you measure it and outcomes. This is where the book gets thought-provoking. Right after the table of contents, Boyle quotes Alan Watts from *Out of Your Mind*, a 2004 audio recording. Watts describes trying to grab a fish. It wiggles and slips from our grasp, so we throw a net over it. “That’s our way of measuring the world—nets with so many holes up and down . . .” It’s a rambling quote, but I think it outlines the problem of measurement. It’s axiomatic that to manage something you have to measure it, or how else do you know if what you are doing has the effect you want? Lots of what we care about—happiness, success, good health—is important and we know what we’re talking about when we’re talking about it, but what are these things and how do we measure them?

This had me thinking that “operational definitions” as used in psychological research are useful. An operational definition is a description of something in terms of the operations (procedures or measurement tools) by which it can be observed and measured. As an undergraduate, I was part of a research project to see if there was any scientific basis for saying people born under different sun signs had different personalities. We all know that people have personality, but we can argue for a very long time with no conclusion about what personality is exactly and how to describe it. (Is his talking over other people his personality, or a function of his being deaf?) The operational definition our team adopted was that personality is what the Omnibus Personality Inventory measures; a person’s scores on the 14 scales are a measure of that person’s personality. We could work with this—but the point for my argument here was that we were aware that we had simply chosen a tool for

defining and measuring personality for the purposes of our study. It was a substitute for personality, which was the thing we cared about; the OPI was not the thing itself.

After about 230 pages, Boyle gives twelve steps to fight back against tickbox, steps such as refusing to provide feedback on a scale, refusing to use robots at store checkouts, and refusing to categorize yourself on feedback or monitoring forms. The list of things to do takes less than eight pages.

The important take-away for *Protocol* readers is not the short twelve-step plan, but making us think seriously about what we are trying to manage, how we can try to manage it, and how to assess whether we are actually accomplishing what we want to accomplish. People don't fit into neat categories. Shows are not all the same. Performers work hard to be unique. Checklists are useful tools for helping us not overlook something, but they have to be used by competent people, people not losing sight of the bigger picture, even while dealing with minutia. ■



**Karl G. Ruling** is ESTA's Senior Technical Standards Manager. He also serves as *Protocol's* Technical Editor. Karl can be reached at [karl.ruling@esta.org](mailto:karl.ruling@esta.org).

*Tickbox*, by David Boyle.  
London: Little, Brown, 2020.  
262 pages. Trade paperback  
ISBN-13: 9781408711873, \$13.84.