

The little engine who could not: “rehabilitating” the individual in safety research

Sidney W. A. Dekker · James M. Nyce ·
Douglas J. Myers

Received: 6 March 2012 / Accepted: 21 May 2012 / Published online: 8 June 2012
© Springer-Verlag London Limited 2012

Abstract Safety science is one of the enduring enlightenment projects, which believes that rationality can create a better, more controllable world. Accidents are not seen as meaningless coincidences, but as failures of risk management, as something that can be improved in the future. The tragedy of safety research is that it has to simultaneously deny and affirm the primacy of human agency. As it has gradually expanded away from the sharp end to see accidents as bureaucratic or administrative in origin, the research keeps supplying linguistic and analytic resources that focus on individual shortcomings in leadership, communication or supervision. This paper concludes that individual human agency is useful to safety work, but not just as an instrument of political or organizational expedience. It is useful because it deeply reflects and reinforces how in the West we understand failure and success. The explanatory power of this discourse is confirmed or taken for granted by safety researchers because it appears so ordinary, self-evident and commonsensical.

Keywords Safety science · Risk management · Accident investigation · Human error · Human agency · Enlightenment

S. W. A. Dekker (✉)
School of Humanities, Griffith University,
Nathan, QLD 4111, Australia
e-mail: s.dekker@griffith.edu.au

J. M. Nyce
Department of Anthropology, Ball State University,
Muncie, IN 47306, USA
e-mail: jnyce@rocketmail.com

D. J. Myers
Department of Community and Family Medicine, Duke
University, Duke Box 3834, Durham, NC 27710, USA
e-mail: dougmyers@nc.rr.com

1 The usefulness of human error I

In this journal not long ago, Cook and Nemeth (2010) reflected on the continued usefulness of the term “human error” particularly as explanation for system or safety failures. The notion of “error,” they observed, is useful to practitioners, organizations and institutions because it can finesse social features or manufacturer flaws “that organizations wish to remain ignorant of or hide” (p. 92). The idea that operator error is the result of (elite-driven) political or organizational processes has appeared in the safety literature before (Perrow 1984; Hollnagel 1998), and more distantly located historical evidence of the political, organizational and psychological usefulness of operator error is easy to find. Accounts of elite-engineered witch hunts, for example, appear as early as in medieval Europe where some physicians accused women of witchcraft to deflect popular discontent with their own inabilities to cure or prevent disease (Levack 1987). Cook and Nemeth (2010) summarize the contemporary case up well: “human (operator) error serves as a kind of lightning rod that conducts the potentially harmful consequences produced by an accident along an (organizationally) safe pathway (p. 91).”

It also, they suggested, has an assuaging function. The conclusion that a human error was to blame offers the illusion of control: “Situating error in the individual raises the prospect of creating an orderly, rational world in which accidents are less likely... Organizations and institutions need to assert that they have control over the circumstances that gave rise to a failure in order to retain independent authority and freedom of action and to restore a public image of reliability” (p. 91). Such a world or the imagined control of it is of course highly desirable if only because it helps provide ontological security. Many, Nietzsche among

them, have commented on the psychologically debilitating and practically paralyzing effects that can occur when a cause is not found (Douglas 1992; Galison 2000; Snook 2000; Hollnagel 2004). Epistemological Culture Theory also suggests that the “raison d’être” and cultural devices like “explanation” not only allow us to believe that we understand the world, but that these devices further contribute to our belief that the world is an ordered reality. Being anxious about not having a cause or explanation is worse than being wrong. This is where “human error” does useful work: being fair in attributing cause and responsibility is less important than being appeased by having found at least something (or someone) to point to. One scapegoat is a small price to pay for the illusion that we know how a world full of risk works (Dekker 2005). So coming up with a finding of human error can help enormously reconfirm the order and nature of things (Cook and Nemeth 2010): “If failure comes from individual error, then attention may be safely directed to restrain or contain the individual” (p. 91). The recent capsizing of the cruise ship *Costa Concordia* prompted comparisons with the *Titanic* disaster and generated such reflections: “In truth it was bad engineering on the *Titanic* that caused its problems. And it was an individual case of human carelessness that caused the *Concordia* to sink. Both of these things can be easily overcome” (O’Neill 2012). For a manager or regulator such a conclusion might suggest that additional rules or procedures, demotions, more training or tighter supervision might take care of the problem (Dekker 2003). And if it does not, then criminal prosecution might (Dekker 2011a, b).

2 The usefulness of human error II

The emerging consensus over the last 40 years points to distal and diffuse factors and explanations instead. Accidents and disasters are seen as organizational and administrative in nature. They are the end-result of an incubation of latent errors and events that are at odds with what is culturally taken for granted but are bureaucratically discounted or ignored (Turner 1978). Accidents, according to the emerging orthodoxy, point to organizations, institutions, to cultures and political environments—not individuals. Thus, the science has increasingly sought the seeds of failure in organizations, in administrations, in corporate cultures and away from the assessments and actions of individual actors. The same organizational, cultural, administrative and bureaucratic processes and arrangements that normally assure successful outcomes can contribute to the build-up of hazard. The focus has thus shifted to upstream or blunt-end factors, to organizational,

operational and other weaknesses and problems that, if left in place, will continue to pose safety risks—independent of who is operating the system at the sharp end. This position has been embraced as the basis of many of today’s investigative agencies and practices.

But of course, if accidents can be investigated, it means that they are not just meaningless coincidences of space and time but rather failures of organizational risk management (Green 2003). Accident investigation, and the safety science on which it is founded, constitutes a project that seems to strongly reflect the enlightenment ideal to be both rational and pragmatic. Accidents not only indicate that something was broken. They also represent something that can be fixed. They express a triumph of reason, an ideological commitment to the position that risk can be better managed in the future. Science, the highest expression of reason, can help people understand the world, control the world, make the world a better place. This has had a paradoxical effect. Safety research has to simultaneously deny and affirm the primacy of human agency. Despite its increasing focus on the blunt end, human agency, even when deeply buried in organizations or cultures, is at the heart of the safety project. At the end of the forensic trails that push deeply into administrations and bureaucracies and cultures, the science still typically finds individuals. It finds people who, despite their efforts, have failed to do their part in managing risk adequately—however far removed, they might have been from the sharp end and the time and place of the accident. As confirmed recently, (Weick and Sutcliffe 2007):

...failure means that there was a lapse in detection. Someone somewhere didn’t anticipate what and how things could go wrong. Something was not caught as soon as it could have been caught (p. 93).

This constant gravitational pull to human agency has been evident from the very earliest days of organizational safety thinking. Categories such as “deficient supervision” and “inadequate leadership” were introduced as part of the early system-thinker’s lexicon (Reason 1990). This has since been extended. Even when probing the blunt end of organizations, safety classification schemes now contain blunt-end failures to provide guidance or oversight, failures to provide training or correct data or other evidence of unsafe leadership and supervision (Shappell and Wiegmann 2001). Human agency is still, or again, seen to be the logical endpoint. In fact, the entire field of ergonomics or human factors has never strayed very far from human agency as primary “lever.” This is reproduced and reconfirmed in the language and models of the field itself, which supplies the linguistic resources (violations, errors, non-compliance) to keep focusing on humans and what they do or do not do.

As Cook and Nemeth have indicated, the objection to diffuse causes and responsibilities is, at least superficially, pragmatic. Referring to the ever-widening search for remote causes of system failure, Reason asked in 1999 “whether this process has gone too far” (p. 217) and concluded that perhaps we should revisit the individual. Citing practical reasons—most remote factors are beyond the reach of system managers—he conceded that with the ever-widening search for upstream factors, the pendulum may have swung too far in the wrong direction. The further away from the sharp end, the more quickly returns on safety will diminish. The distribution of articles in, for example, the patient safety literature seems to agree. Few recent studies into healthcare risk and safety actively looked at teams or organizations, while a review of 360 such journal articles found at least 98 focused on individual as the site for intervention and improvement (Waterson 2009). Speaking at a conference of air safety investigators, Young et al. (2004) cited various examples where the official investigations, in keeping with safety-scientific orthodoxy, found large numbers of remote factors. But, they asked, what about the “incorrectly handled situations” and “irrational acts” by front-end operators which (seemingly) made all the difference? One of these examples was a Boeing 747 airliner descending into a wet and rainy Bangkok. The Captain instructed the First Officer to break off the approach and go around. Four seconds later, however, the Captain took control of the aircraft, retarded the thrust levers to land the jet. It overran the runway and sustained substantial damage. To Young and colleagues, this accident could not be explained with reference to deficiencies in company procedures or insufficient training for landing on waterlogged runways and other latent factors. It made sense to them only if we focused on the sharp end of practice: the “active and irrational error” by the Captain (p. 4). Citing the overwhelming role of errors of commission at the sharp end in most accidents they go on to critique the “dogmatic” application of safety science’s own new models—ones that focus on remote, blunt-end factors and the latent conditions in them (Young et al. 2004):

One of the main implications of this has been the tenacious and dogmatic search for latent conditions leading up to the incident. Overzealous implementation...has led to an illusion of management responsibility for *all* errors. While this may very often be the case for major accidents, in other cases [it] seems contrived and untenable. (p. 1) (emphasis in original).

Such research uncritically collapses social science with moral philosophy. This is a marriage which even Durkheim would have warned us against. It reduces (or confuses) social analysis with moral accountancy, something that represents popular belief more than the outcome of any empirically based investigation or science.

An analysis of the NASA Shuttle Challenger and Columbia accidents by Feldman epitomizes this confusion (Feldman 2004). At first, the analysis and its findings remain consistent with the standard safety-scientific model: Misunderstandings of flight risk were part and parcel of how the system was rigged—bureaucratically, politically, professionally. They were the systemic products of overconfidence in quantitative data, a marginalization of non-quantitative data, an insensitivity to uncertainty and loss of organization memory, of the illusion that engineering problems and solutions can be addressed independently from organizational goals. All this led to inaccurate estimations of risk. Vaughan (1996, 2005) similarly showed how organizational culture, structural bureaucratic secrets, and the gradual normalization of deviant launch results meant that engineers and managers legitimately and explicitly accepted ever greater risk. The Columbia accident investigation itself even concluded that organizational cultures, post-cold war politics and organizational history were among the accident causes (CAIB 2003). Together these things leave very little room, if any, for human agency. In his analysis of the Challenger launch decision, Jensen (1996) argued that the effects of such decision-making contexts can be so powerful, that:

We should not expect the experts to intervene, nor should we believe that they always know what they are doing. Often they have no idea, having been blinded to the situation in which they are involved. These days, it is not unusual for engineers and scientists working within systems to be so specialized that they have long given up trying to understand the system as a whole, with all its technical, political, financial and social aspects (p. 368).

Yet, despite his own analysis, this is precisely Feldman’s solution. In his conclusion, he exhorts engineers to intervene, to know what they are doing, to speak up, to not be blinded to the situation(s) in which they are involved. “Engineers need intense cultivation of their professional responsibilities within organizations,” he argued (p. 713). In other words, the way to deal with threat and evil structures (Zimbardo 2008) is for individuals to be more Protestant, that is, to work harder, to be more conscientious and virtuous. Feldman continued:

Engineering societies need to require engineers to act in accordance with the prevent-harm ethic. This requirement must include both training to inculcate the prevent-harm ethic and sanctions—up to losing one’s license—when the ethic is violated (p. 714).

This rallying cry would be well-received in medicine. The medical profession, ever committed to its simultaneous but contradictory belief in individual autonomy and denial of

individual fallibility, tends to argue that good outcomes are the result of “good doctoring,” of competent people who can succeed often in spite of the organization. Pellegrino (2004) contended that rather than “systems,” healthcare needs individuals with “strength of character to be virtuous” (p. 90). This defense of individual autonomy reflects the view that medical training is the kind of moral enterprise which, if left alone, can sort the sheep from the goats, the ethical from the morally deficient (Bosk 2003). Promoting “systems thinking,” Pellegrino argued, undermines the unique fiduciary relationship between caregiver and patient and shortcuts personal control over, and accountability for, clinical outcomes. The belief that risk is something embedded in (and best addressed in terms of) the personal relationship between patient and caregiver still dominates in modern healthcare despite much evidence to the contrary (DeVille 2004). As Gawande (2002) admitted: “Many doctors don’t like to talk of ‘systems problems.’ Something in me, too, demands an acknowledgment of my autonomy...my ultimate culpability” (p. 73). What Bosk found, after studying surgeons at a teaching hospital, was that surgeons considered virtually all “technical” errors remediable and things that more skills training and practice could mostly solve. Surgeons had much more trouble, however, with what they considered “normative” errors. These, Bosk observed, were not errors *in* a role, like the technical errors, but errors *of* a role, failures to live up to the moral duty of what it meant to be a physician. In particular, these were failures to live up to the expectation to act both responsibly and autonomously. Surgical trainees who continued to make such normative errors were almost always terminated (Bosk 2003).

As Feldman suggested, the unique fiduciary relationship in which the expert understands something that others do not, places a higher moral burden on that expert—the burden to speak up, to act in accordance with the harm prevention ethic (as medicine has tried since Hippocrates). Though Feldman took things a bit further (with sanctions for violations), his appeal is consistent with calls for the empowerment and legitimation of organizationally less powerful groups. Such strategies have become popular in aviation, healthcare, shipping as well as other industries (under names such as Crew Resource Management training). The empowerment of nurses, another example, is seen as key to solving many patient safety problems by some (Pronovost and Vohr 2010). The question is whether such a Thomist moral script can have the kind of effect its advocates claim for it (Salas et al. 2006). In the end, this appeal to the individual may be nothing more than a kind of Trojan horse that conceals and finesses where power and responsibility lie. The problem is that the position Feldman took can actually reinforce the role that hierarchy and elites play in today’s organizations and their “restorative”

enterprises. Seemingly emancipatory strategies often take for granted who defines “below” and “above” and thus ally themselves with who sets up the game and establishes its rules—as well as who largely “wins.” Feldman discounted the role relations of dependency, particularly those that suppress self-revelation and staff communication, as well as all the other “ethical” behaviors that Feldman seemed to think all organizational actors possess and can exercise. Unfortunately, no questions have yet been asked in the safety literature about the often considerable resources elites have to co-opt or subvert such empowerment or emancipatory initiatives. Strategies that would empower the individual, like those that focus on “just cultures” in hospitals, for example (Marx 2001), seem only to continue or reinforce management control over staff. Most safety science and “just culture” advocates tend to focus on the formal, quantifiable aspects of the organization and its actors. The numerous attempts to measure both organization culture and just culture(s) represent one example of this. The irony is that the attention given to these dimensions of organizational life can deflect analytic attention from where power lies and how it is exercised in the modern organization (Sagan 1994).

3 Discussion and conclusion

Why has the fundamental dilemma of safety work become more explicit in the literature and the discourse of the field? The last decades in the West have seen a gradual reduction in the acceptance of risk altogether (Beck 1992), and the expectation that some safety-critical activities are accident-free, with a zero-tolerance of failure. The increasingly flawless performance of some systems has led to a belief in infallibility, beyond these particular systems, and thus to a growing intolerance of failure (Amalberti 2001). Experts are expected to make any residual accidents comprehensible, which often means explaining which risk factors were not controlled by whom. This is often easier to define and assess in terms of individual choices about procedure and action. Political economists might point out that the gradual erosion and loss of union (and thus worker) power since the 1980s might be a factor here too, as well as deregulation across industries. The latter has been linked to the rise of “culture” as explanation for both organizational success and failure (Silbey 2009). There have also been fundamental changes in the relationship between government and corporations. Managing liabilities in the wake of accidents has become more important. As Cook and Nemeth pointed out, this in turn can lead to a search for scapegoats or targets that can deflect any higher-order or organizational responsibility for failure. The media also doubtlessly has celebrated certain accidents while

deflecting or ignoring others. Media coverage of negative events can articulate and animate social reactions by constructing anti-heroes (Elkin 1955; McLean and Elkind 2004). The result could be seen as a strong democratic project (Foucault 1975; Anon 2005) where the polity, the elite, responds to and “fairly” represents the concerns of the society in which it operates. Yet in the last decades, there has been an emergence of an electronically mediated democratization and the increasing accessibility of knowledge, as well as wider consumer vocalism and activism. These can put failings of complex systems (or alleged failings of individuals in them) on fuller display than before (Anon 2005), independent of mainstream media. Also within modern society itself, there seems to be a desire for morally and epistemologically unambiguous, that is, actor-focused explanations for bad events (Giddens 1991). And there is, of course, the change in the academic enterprise itself. There has been a (re)turn to a more instrumental kinds of explanations. The boundaries drawn between corporations and universities have become less clear and, accordingly, the merger between corporate and research agendas may have reinforced a tendency to look for instrumental and “simple” explanations for complex events.

Above all, it is possible that the move away from human error, from the individual, was always tentative. After, all, the idea of individual error is useful not just as instrument of political or organizational expedience. It is useful because it deeply reflects and reinforces how in the West we understand failure and success. The explanatory power of this discourse is confirmed (or taken for granted) by safety researchers because it appears so ordinary, self-evident and commonsensical. Perhaps it is not so unexpected that even safety researchers would return to the individual and individual responsibility. The thrust of safety research toward systems and remote factors and diffuse responsibilities seems to run counter to “common sense” and science itself. Moreover, safety researchers are informed by—and also a formative part of—the Western moral enterprise which elevates responsibility, choice and autonomy above almost any considerations, a stance reinforced by Judeo-Christianity. After all, whom did Job blame for the injuries inflicted upon him—God or himself? It can be argued that the discourse in the science of safety, with its focus on human agency wherever located in system hierarchies, provides intellectual resources (among them languages and models) that reconfirm and reinforce Western notions of individualism, rationality, morality and accountability (Dekker 2011a, b). This in turn defines (and circumscribes) the repertoire of solutions and countermeasures available to the safety science community. To frame explanation and intervention in reference to system and structure, not the individual,

raises the specter of whether this community can “really” do science when it so ignores what Western culture insists is common sense.

If it could, then a research agenda that starts to pull away from this heritage would have to accomplish multiple things. First, it is clear that safety science and cognitive engineering need to continue investing in projects that trace the intellectual history, or histories, of the field. This is not so much to provide a genealogy, but rather to perform a kind of archeology that looks at what we in the community have taken for granted, and to bring to light the institutional and cultural origins and influences behind our assumptions. This calls for collective self-reflection, to correct a kind of myopia that every intellectual community suffers from but seldom attempts to make visible. The second step would be to apply and test what we have learned from this to pieces of analytic and empirical research. Both these steps would have analytic and practical yield.

References

- Amalberti R (2001) The paradoxes of almost totally safe transportation systems. *Saf Sci* 37(2–3):109–126
- Anon (2005) Murder! Mayhem! Social Order! *Wilson Quarterly* vol 29. Woodrow Wilson International Center for Scholars, Washington, pp 94–96
- Beck U (1992) *Risk society: towards a new modernity*. Sage Publications Ltd, London
- Bosk C (2003) *Forgive and remember: managing medical failure*. University of Chicago Press, Chicago
- CAIB (2003) Report volume 1, August 2003. Columbia Accident Investigation Board, Washington, DC
- Cook RI, Nemeth CP (2010) “Those found responsible have been sacked”: some observations on the usefulness of error. *Cognit Technol Work* 12:87–93
- Dekker SWA (2003) Illusions of explanation: a critical essay on error classification. *Int J Aviat Psychol* 13(2):95–107
- Dekker SWA (2005) Ten questions about human error: a new view of human factors and system safety. Lawrence Erlbaum Associates, Mahwah
- Dekker SWA (2011a) The criminalization of human error in aviation and healthcare: a review. *Saf Sci* 49(2):121–127
- Dekker SWA (2011b) What is rational about killing a patient with an overdose? Enlightenment, continental philosophy and the role of the human subject in system failure. *Ergonomics* 54(8): 679–683
- DeVillie K (2004) God, science, and history: the cultural origins of medical error. In: Sharpe VA (ed) *Accountability: patient safety and policy reform*. Georgetown University Press, Washington, pp 143–158
- Douglas M (1992) *Risk and blame: essays in cultural theory*. Routledge, London
- Elkin F (1955) Hero symbols and audience gratifications. *J Educ Sociol* 29(3):97–107
- Feldman SP (2004) The culture of objectivity: quantification, uncertainty, and the evaluation of risk at NASA. *Human Relat* 57(6):691–718
- Foucault M (1975) *The spectacle of the scaffold*. Penguin Group, London

- Galison P (2000) An accident of history. In: Galison P, Roland A (eds) *Atmospheric flight in the twentieth century*. Kluwer, Dordrecht, The Netherlands pp 3–44
- Gawande A (2002) *Complications: a surgeon's notes on an imperfect science*. Picado, New York
- Giddens A (1991) *Modernity and self-identity: self and society in the late modern age*. Stanford University Press, Stanford
- Green J (2003) The ultimate challenge for risk technologies: controlling the accidental. In: Summerton J, Berner B (eds) *Constructing risk and safety in technological practice*. Routledge, London
- Hollnagel E (1998) *Cognitive reliability and error analysis method: CREAM*. Elsevier, Oxford
- Hollnagel E (2004) *Barriers and accident prevention*. Aldershot, UK
- Jensen C (1996) No downlink: a dramatic narrative about the challenger accident and our time. Farrar, Straus, Giroux, New York
- Levack BP (1987) *The witch-hunt in early modern Europe*. Longman, London
- Marx D (2001) Patient safety and the “just culture”: a primer for health care executives. Columbia University, New York
- McLean B, Elkind P (2004) *The smartest guys in the room: the amazing rise and scandalous fall of Enron*. Portfolio, New York
- O'Neill B (2012) Concordia provides no morality tale. *The Weekend Australian*, Sydney, p 15
- Pellegrino ED (2004) Prevention of medical error: where professional and organizational ethics meet. In: Sharpe VA (ed) *Accountability: patient safety and policy reform*. Georgetown University Press, Washington, pp 83–98
- Perrow C (1984) *Normal accidents: Living with high-risk technologies*. Basic Books, New York
- Pronovost PJ, Vohr E (2010) *Safe patients, smart hospitals*. Hudson Street Press, New York
- Reason JT (1990) *Human error*. Cambridge University Press, New York
- Sagan SD (1994) Toward a political theory of organizational reliability. *J Contingencies Crisis Manag* 2(4):228–240
- Salas E, Wilson KA et al (2006) Does crew resource management training work? An update, an extension, and some critical needs. *Hum Factors* 48(2):392–413
- Shappell SA, Wiegmann DA (2001) Applying reason: the human factors analysis and classification system. *Human Factors Aerosp Saf* 1:59–86
- Silbey S (2009) Taming Prometheus: talk about safety and culture. *Annu Rev of Sociol* 35: 341–369
- Snook SA (2000) *Friendly fire: the accidental shootdown of US Black Hawks over Northern Iraq*. Princeton University Press, Princeton
- Turner BA (1978) *Man-made disasters*. Wykeham Publications, London
- Vaughan D (1996) *The Challenger launch decision: risky technology, culture, and deviance at NASA*. University of Chicago Press, Chicago
- Vaughan D (2005) System effects: on slippery slopes, repeating negative patterns, and learning from mistake? In: Starbuck WH, Farjoun M (eds) *Organization at the limit: lessons from the Columbia disaster*. Blackwell Publishing, Malden, pp 41–59
- Waterson P (2009) A critical review of the systems approach within patient safety research. *Ergonomics* 52(10):1185–1195
- Weick KE, Sutcliffe KM (2007) *Managing the unexpected: Resilient performance in an age of uncertainty*. Jossey-Bass, San Francisco
- Young MS, Shorrock ST et al (2004) Who moved my (Swiss) cheese? The (r)evolution of human factors in transport safety investigation. *International Society of Air Safety Investigators (ISASI)*, Gold Coast
- Zimbardo P (2008) *The Lucifer effect: understanding how good people turn evil*. Random House, New York