

When human error becomes a crime

Sidney W. A. Dekker
Lund University School of Aviation, Sweden

Abstract. Progress on safety means embracing a systems view and moving beyond blame. But recent (as well as historical) incidents and accidents often show the opposite response. Human error sometimes gets turned into a crime by holding the operator (pilot, controller, maintenance technician) criminally liable for the bad outcome. Criminalizing human error is hostile to both the systems view and moving beyond blame. Criminalization singles out individuals under the banner of “holding them accountable” for outcomes that in reality rely on many more contributions. Criminalization also countervenes progress on safety in other important ways. Demonstrated effects include the silencing of incident reporting, polarization of stakeholder attitudes and positions, and destabilization of industrial relations. There is no evidence of any safety benefits. This paper explores some motives behind the criminalization of error and suggests the aviation industry should redesign accountability relationships between its stakeholders.

The criminalization of error

In the aftermath of several recent accidents and incidents (e.g. Wilkinson, 1994; Ballantyne, 2002; Ruitenbergh, 2002) the pilots or air traffic controllers involved were charged with criminal offenses (e.g. professional negligence; manslaughter). Criminal charges differ from civil lawsuits in many respects. Most obviously, the target is not an organization but individuals (air traffic controller(s), flightcrew, maintenance technicians). Punishment consists of possible incarceration or some putatively rehabilitative alternative—not (just) financial compensation. Unlike organizations covered against civil suits, few operators themselves have insurance to pay for legal defense against criminal charges.

Some maintain that criminally pursuing operators for erring on the job is morally non-problematic. The greater good befalls the greater number of people (i.e. all potential passengers) by protecting them from egregiously unreliable operators. A lot of people win, only a few outcasts lose. To human factors, however, this may be utilitarianism inverted. Everybody loses when human error gets criminalized:

Upon the threat of criminal charges, operators stop sending in safety-related information; incident-reporting grinds to a halt (Ruitenbergh, 2002; North, 2002).

Criminal charges against individual operators can polarize industrial relations. If the organization wants to limit civil liability, then official blame on the operator could deflect attention from upstream organizational issues related to training, management, supervision, design decisions. Blaming such organizational issues, in contrast, can be a powerful ingredient in an individual operator’s criminal defense—certainly when the organization has already rendered the operator expendable by euphemism (stand-by, ground duty,

administrative leave) and without legitimate hope of meaningful re-employment. In both cases, industrial relations are destabilized.

Incarceration or alternative punishment of pilots or controllers has no demonstrable rehabilitative effect (perhaps because there is nothing to rehabilitate). It does not make a pilot or air traffic controller (or his or her colleagues) any safer—indeed, the very idea that vicarious learning or redemption is possible through criminal justice is universally controversial.

The aviation community itself shows ambiguity to the criminalization of error. Responding to the 1996 ValuJet accident, where mechanics loaded oxygen generators into the cargo hold of a DC-9 which subsequently caught fire, the editor of *Aviation Week and Space Technology* "strongly believed the failure of SabreTech employees to put caps on oxygen generators constituted willful negligence that led to the killing of 110 passengers and crew. Prosecutors were right to bring chargers. There has to be some fear that not doing one's job correctly could lead to prosecution" (North, 2000, p. 66). Rescinding this two years later, however, the editor opined how learning from accidents and criminal prosecution go together like "oil and water, cats and dogs"; that "criminal probes do not mix well with aviation accident inquiries" (North, 2002, p. 70). Other contributions (e.g. the "No Concord" editorial in *Flight International* 22-28 January 2002) reveal similar instability with regard to prosecuting operators for error. Culpability in aviation does not appear to be a fixed notion, connected unequivocally to features of some incident or accident. Rather, culpability is a highly flexible category—it is negotiable; subject to national and professional interpretations, influenced by political imperatives and organizational pressures, and part of personal or institutional histories.

Punishment or learning?

Human factors (e.g. Reason, 1997; AMA, 1998; Palmer et al., 2001; Woods & Cook, 2002) agrees that progress on safety depends in large part on:

- Taking a systems perspective: Accidents are not caused by failures of individuals, but result from the conflux or alignment of multiple contributory system factors, each necessary and only jointly sufficient. The source of accidents is the system, not its component parts.
- Moving beyond blame: Blame focuses on the supposed defects of individual (front-line) operators and denies the import of systemic contributions. Blame leads to defensive posturing, obfuscation of information, protectionism, polarization, and mute reporting systems.

The same sources also take progress on safety as synonymous to learning from failure. This makes punishment and learning two mutually exclusive activities: You can either learn from an accident or punish the individuals involved in it, but it is probably very difficult to do both at the same time (Dekker, 2002):

- Punishment of individuals protects false beliefs about basically safe systems where humans are the least reliable components. Learning challenges and potentially changes our beliefs about what creates safety.

- Punishment emphasizes that failures are deviant, that they do not naturally belong in the system. Learning means that failures are seen as “normal”—as resulting from the inherent pursuit of success in resource-constrained, uncertain environments.
- Punishment turns the culprits into unique and necessary ingredients for the failure to happen. Learning means that every operator and operation is potentially vulnerable to breakdown.
- Punishment compartmentalizes safety lessons (this person, that department). Learning generalizes safety lessons for everybody’s benefit.
- Punishment conditions others to not get caught next time. Learning is about avoiding a next time altogether
- Punishment is about the search for closure, about moving beyond and away from the terrible event. Learning is about continuous improvement, about closely integrating the event in what the system knows about itself.

While research on the hindsight bias (e.g. Fischhoff, 1975) and the fundamental attribution error (e.g. Cheng & Novick, 1992) helps explain the mechanisms by which we blame front-line operators and ignore other contributions to failure, it sheds little light on why we continue to do so (for example through criminal prosecution) in the face of overwhelming evidence that (1) individual failure is not what causes accidents, and (2) blame and punishment are antithetical to making progress on safety (e.g. Maurino et al., 1995; Woods & Cook, 2002). Other, deeper biases or reflexes in our reactions to failure must be at work to conspire against fully embracing a systems perspective—forces that we would have to understand in order to make our chosen route to progress on safety really successful. In this paper I explore contributions from a brief range of disciplines potentially able to shed some more light on the issue.

Social cognition research on accountability

One foundational contribution to our understanding of blame and punishment comes from social cognition research on accountability. Accountability is fundamental to any social relation because of the implicit or explicit expectation that one may be called upon to justify one’s beliefs and actions to others. The social functionalist argument for accountability is that this expectation is mutual: as social beings we are locked into reciprocating relationships. Accountability, however, is not a unitary concept—even if this is what many stakeholders may think when aiming to improve people’s performance under the banner of “holding them accountable”. There are as many types of accountability as there are distinct relationships among people, and between people and organizations, and only highly specialized subtypes of accountability actually compell people to expend more cognitive effort. Expending greater effort, moreover, does not necessarily mean better task performance, as operators may become concerned more with limiting exposure and liability than with performing well (Lerner & Tetlock, 1999), something that can be observed in the decline of incident reporting with threats of prosecution (North, 2002). What is more, if accounting is perceived as illegitimate, for example intrusive, insulting or ignorant of real work, then any beneficial effects of accountability will vanish or backfire. Effects that have been experimentally demonstrated include a decline in motivation, excessive stress and attitude polarization (Lerner & Tetlock, 1999). These effects indeed occurred in cases (e.g. Ballantyne, 2002; Ruitenbergh, 2002) where pilots and air traffic controllers were “held

accountable” by courts and constituencies unaware of the real trade-offs and dilemmas that make up actual operational work (see Woods & Cook, 2002).

The research base on social cognition, then, tells us that accountability, even if inherent in human relationships, is not unambiguous or non-problematic. The good side of this is that if accountability can take many forms, then alternative, perhaps more productive avenues of holding people “accountable” are possible. Giving an account, after all, does not have to mean exposing oneself to liability, but rather “telling one’s story” so that others can learn vicariously. Many sources, even within human factors, point to the value of storytelling in preparing operators for complex, dynamic situations in which not everything can be anticipated (e.g. AMA, 1998; Klein, 1998). Stories are easily remembered, scenario-based plots with actors, intentions, clues and outcomes that in one way or another can be mapped onto current difficult situations and matched for possible ways out. Incident reporting systems can capitalize on this possibility, while more incriminating forms of “accountability” actually retard this very quality by robbing from people the incentive to tell stories in the first place.

The anthropological approach

The anthropologist is not so intrigued by flaws in people’s reasoning process that produce for example the hindsight bias, but wants to know something about casting blame. Why is blame a meaningful response for those doing the blaming? Peoples are organized in part by the way in which they explain misfortune and subsequently pursue retribution or dispense justice. Societies tend to rely on one dominant model of possible cause from which they construct a plausible explanation (Douglas, 1992). For example, in the moralistic model, misfortune results from offending ancestors, sinning, or breaking some taboo. Each explanation is followed by a fixed repertoire of obligatory actions that follow on that choice. If taboos are broken, for example, then rehabilitation is demanded through expiatory actions (garnering forgiveness through some purification ritual), evidence of which can be seen in the aftermath of at least one recent accident (Ballantyne, 2002).

In the extrogenous model, external enemies of the system are to blame for misfortune—a response that can be observed even today in the demotion or exile of “failed” operators: pilots or controllers or technicians who, ex-post-facto, are relegated to a kind of underclass that no longer represents the professional corps. The ritualistic reappropriation of badges, certificates, stripes, licences, uniforms, or other identity and status markings in the wake of an accident (e.g. Wilkinson, 1994; Ballantyne, 2002) delegitimizes the errant operator as member of the operational community. A part of such derogation, of course, is psychological defense on the part of (former) colleagues who would need to distance themselves from a realization of equal vulnerability to similar failures. Yet such delegitimization also makes criminalization easier by beginning the incremental process of dehumanizing the operator in question. Wilkinson (1994) presents an excellent example of such demonification in the consequences that befell a British Airways B-747 pilot after allegedly narrowly missing a hotel at Heathrow in thick fog. Demonification there proved incremental in the sense that it not only made criminal pursuit possible in the first place, but subsequently necessary. It fed on itself: demons such as this pilot would need to be punished, demoted, exorcised. The tabloid press had a large share in dramatizing the case, promoting the captain’s dehumanization to the point where his suicide was the only way out.

Failure and fear

Today, almost every misfortune is followed by questions centering on “whose fault?”, and “what damages, compensation?” Every death must be chargeable to somebody’s account. Such responses approximate the primitives’ resistance to the idea of natural death remarkably well (Douglas, 1992). Death, even today, is not considered natural—it has to result from starker causes. Such resistance to the notion that deaths actually can be “accidental” is obvious in responses to recent mishaps. For example, Snook (2000, p. 203) comments on his own disbelief, his struggle, in analyzing the friendly shootdown of two US Black Hawk helicopters by US Fighter Jets over Northern Iraq in 1993:

"This journey played with my emotions. When I first examined the data, I went in puzzled, angry, and disappointed—puzzled how two highly trained Air Force pilots could make such a deadly mistake; angry at how an entire crew of AWACS controllers could sit by and watch a tragedy develop without taking action; and disappointed at how dysfunctional Task Force OPC must have been to have not better integrated helicopters into its air operations. Each time I went in hot and suspicious. Each time I came out sympathetic and unnerved... If no one did anything wrong; if there were no unexplainable surprises at any level of analysis; if nothing was abnormal from a behavioral and organizational perspective; then what have we learned?"

Snook confronts the question of whether learning, or any kind of progress on safety, is possible at all if we can find no wrongdoing, no surprises; if we cannot find some kind of deviance. If everything was normal, then how could the system fail (cf. Perrow, 1984)? Indeed, this must be among the greater fears that define Western society today. Investigations that do not turn up a “Eureka part”, as the label became in the TWA800 probe, are feared not because they are bad investigations, but because they are scary. No Eureka part, no fault nucleus, no seed of destruction. Rather, failure results from doing business as usual (Perrow, 1984; Dekker, 2002). As Galison notes (2000, p. 32):

"If there is no seed, if the bramble of cause, agency, and procedure does not issue from a fault nucleus, but is rather unstably perched between scales, between human and non-human, and between protocol and judgment, then the world is a more disordered and dangerous place. Accident reports, and much of the history we write, struggle, incompletely and unstably, to hold that nightmare at bay."

Galison’s account is evidence of this fear (this “nightmare”) of not being in control over the systems we design, build and operate, where failures emanate from normal everyday interactions rather than from traceable, controllable single seeds or nuclei.

Being afraid may be worse than being wrong. Selecting a scapegoat to carry the interpretive load of an accident or incident is the “easy” price we pay for our illusion that we actually have total control over our risky technologies (Perrow, 1984; Pagels, 1988). Sending controllers or pilots or maintenance technicians to jail may be morally wrenching (but not unequivocally so — remember North, 2000), but it is preferable over its scary alternative: acknowledging that we do not have total control over the risky technologies we build and consume. The alternative would force us to admit that failure is an emergent

property, that "mistake, mishap and disaster are socially organized and systematically produced by social structures"; that these mistakes are normal; to be expected because they are "embedded in the banality of organizational life" (Vaughan, 1996, p. xiv). It would force us to acknowledge the relentless inevitability of mistake in organizations; to see that harmful outcomes can occur in the organizations constructed to prevent them; that harmful consequences can occur even when everybody follows the rules (Vaughan, 1996).

Preferring to be wrong over being afraid would also be consistent with the common reflex towards individual responsibility in the West. Why, indeed, are the hindsight bias and fundamental attribution error—both of which train our focus onto an individual's misassessments and wrong actions as precursors to failure—prerational in our understanding of misfortune? St. Augustine, the deeply influential moral thinker in Judeo-Christian society, saw human suffering as occurring not only because of individual human fault (Pagels, 1988), but because of human choice — the conscious, deliberate, rational choice to err. The idea of a rational choice to err is pervasive in Western thinking (cf. Reason, 1997), almost to the point of going unnoticed, unquestioned, because it makes such "common sense". The idea is that pilots have a choice to take the correct runway but fail to take it (i.e. they make the wrong choice because of attentional deficiencies or motivational shortcomings, despite the cues that were available and time they had to evaluate those cues) (Ballantyne, 2002); controllers have a choice to see a looming conflict, but elect to pay no attention to it (Ruitenbergh, 2002). After the fact, it often seems as if people "chose" to err, despite all available evidence indicating they had it wrong.

The story of Adam's original sin, and especially what St. Augustine makes of it, reveals the same space for conscious negotiation that we retrospectively invoke on behalf of people carrying out safety-critical work in real conditions (Eve has a deliberative conversation with the snake on whether to sin or not to sin; on whether to err or not to err). It emphasizes the same conscious presence of cues and incentives to not err — yet Adam elects to err anyway. The prototypical story of error/violation and its consequences in Judeo-Christian tradition tells of people who are equipped with the requisite intellect, who have received the appropriate indoctrination (don't eat that fruit), who display capacity for reflective judgment and who actually have the time to choose between a right and a wrong alternative. They then proceed to pick the wrong alternative — a choice that would make a big difference for their lives and the lives of others. It is likely that rather than "causing" the Fall into continued error, as St. Augustine would have it, Adam's original sin portrays how we think about error, and how we have thought about it for ages. The idea of free will permeates our moral thinking, and most probably influences how we look at human performance to this day.

Of course this "illusion of free will" (Reason, 1997), though dominant in post-hoc analyses of error, is at odds with the real conditions under which people perform work: where resource limitations and uncertainty severely constrain the choices open to them. Van den Hoven (1999) calls this "the pressure condition". Operators such as pilots and air traffic controllers are "narrowly embedded"; they are "configured in an environment and assigned a place which will provide them with observational or derived knowledge of relevant facts and states of affairs" (p. 3). Such environments are exceedingly hostile to the kind of reflection necessary to meet the regulative ideal of individual moral responsibility — the kind espoused in the story of Adam and Eve and the kind retrospectively presumed on behalf of operators in difficult situations that led to an accident or incident.

Human factors refers to this as an authority-responsibility double bind (see Woods & Cook, 2002; Dekker, 2002): a mismatch occurs between the responsibility expected of

people to do the right thing, and the authority given or available to them to live up to that responsibility. Society expresses its confidence in operators' responsibility through payments, status, symbols and the like. Yet operators' authority may fall short of that responsibility in many important ways. Operators typically do not have the degrees of freedom assumed by their professional responsibility because of a variety of reasons: practice is driven by multiple goals that may be incompatible (simultaneously having to achieve maximum capacity utilization, economic aims, customer service, safety) (Woods & Cook, 1992). As Wilkinson (1994, p. 87) remarks: "A lot of lip service is paid to the myth of command residing in the cockpit, to the fantasy of the captain as ultimate decision-maker. But today the commander must first consult with the accountant".

Authority to make the decisions for which you are responsible is not only constrained by goals other than safety. In addition, authority can be limited because time and other resources for making sense of a situation are lacking (Van den Hoven, 1999); information may not be at hand or may be ambiguous (e.g. Klein, 1998); and there may be no neutral or additional expertise to draw on. Only recent additions to the human factors literature (e.g. in the form of ecological task analysis, see Vicente, 1999; Flach, 2000) explicitly take these and other constraints on people's practice into consideration in the design and understanding of work.

Free will is a logical impossibility in cases where there is a mismatch between responsibility and authority. Which is to say that free will is always a logical impossibility in real work domains where real safety-critical work is carried out. When the notion or illusion of free will disappears, then with it disappears the basis for traditional Judeo-Christian notions of responsibility and culpability. As said earlier, it is time to re-assess and re-invent the ways in which the aviation industry wants to hold its own "accountable".

Blame-free cultures?

Blame free cultures are extremely rare. Examples have been found among Sherpas in Nepal (Douglas, 1992), who pressure each other to settle quarrels peacefully and reduce rivalries with strong informal procedures for reconciliation. Laying blame accurately is considered much less important than a generous treatment of the victim. Sherpas irrigate their social system with a lavish flow of gifts, taxing themselves collectively to ensure nobody goes neglected, and victims are not left exposed to impoverishment or discrimination (Douglas, 1992). This mirrors the propensity of Scandinavian cultures for collective taxation and thick webs of social security, where indeed (criminal) prosecution or lawsuits in case of accidents are rare. US responses stand in stark contrast (although criminal prosecution of operators is rare there). Despite plentiful litigation, victims are typically undercompensated (Palmer et al. 2001). A blame-free culture then, may depend more on consistently generous treatment of victims than on denying that professional "accountability" exists.

In fact, holding people accountable can be consistent with being blame-free if only we have the creativity to think in novel ways about accountability, which would involve innovations in our relationships among the various stakeholders. Indeed, in order to continue making progress on safety, the aviation industry should reconsider and reconstruct accountability relationships between its stakeholders (organizations, regulators, litigators, operators, passengers). In the adversarial posturing that the criminalization of error generates today, "truth" becomes fragmented across multiple versions that advocate

particular agendas (staying out of jail; limiting corporate liability) making learning from the mishap almost impossible. Instead, operators involved in mishaps could be held "accountable" by inviting them to tell their story (their "account"), by then systematizing and distributing the lessons in it, and by using this to sponsor vicarious learning for all. Perhaps such notions of accountability would be better able to move us in the direction of an as yet elusive blame-free culture.

References

- Ballantyne, T. (2002). Peers support sieged pilots: Taipei accused of glossing over airport deficiencies. *Orient Aviation*, June, pp. 20-21.
- Cheng, P. & Novick, L. (1992). Covariation in natural causal induction. *Psychological Review*, 99(2), 365-382.
- Dekker, S. W. A. (2002). *The field guide to human error investigations*. Bedford, UK: Cranfield University Press/Aldershot, UK: Ashgate.
- Douglas, M. (1992). *Risk and blame: Essays in cultural theory*. London, UK: Routledge.
- Fischhoff, B. (1975). Hindsight is not foresight: The effect of outcome knowledge on judgment under uncertainty. *Journal of Experimental Psychology: Human Perception and Performance*, 1(3), 288-299.
- Flach, J. M. (2000). Discovering situated meaning: An ecological approach to task analysis. In J. Schraagen, S. Chipman, & V. Shalin (Eds.). *Cognitive Task Analysis*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Galison, P. (2000). An accident of history. In P. Galison, & A. Roland (eds.), *Atmospheric flight in the twentieth century*, pp. 3-44. Dordrecht, NL: Kluwer Academic Publishers.
- Hoven, M. J. van den (2001). *Moral responsibility and information technology*. Rotterdam, the Netherlands: Erasmus University Center for Philosophy of ICT.
- Klein, G. (1998). *Sources of power: How people make decisions*. Cambridge, MA: MIT Press.
- Lerner, J. S., & Tetlock, P. E. (1999). Accounting for the effects of accountability. *Psychological Bulletin*, 125(2), 255-275.
- Maurino, D. E., Reaon, J. T., Johnston, N., & Lee, R. B. (1995). *Beyond aviation human factors*. Aldershot, UK: Ashgate Publishing Co.
- North, D. M. (2000). Let judicial system run its course in crash cases. *Aviation Week and Space Technology*, May 15, p. 66.
- North, D. M. (2002). Oil and water, cats and dogs. *Aviation Week and Space Technology*, February 4, p. 70.
- Pagels, E. (1988). *Adam, Eve and the serpent*. London: Weidenfeld & Nicolson.
- Palmer, L. I., Emanuel, L. L., & Woods, D. D. (2001). *Managing systems of accountability for patient safety*. Washington, DC: National Health Care Safety Council of the National Patient Safety Foundation.
- Perrow, C. (1984). *Normal accidents: Living with high-risk technologies*. New York: Basic Books.
- Reason, J. T. (1997). *Managing the risks of organizational accidents*. Aldershot, UK: Ashgate Publishing Co.
- Ruitenbergh, B. (2002). *Court case against Dutch Air Traffic Controllers*. Amsterdam, the Netherlands: LVNL.
- Snook, S. A. (2000). *Friendly fire: The accidental shootdown of US Black Hawks over Northern Iraq*. Princeton, NJ: Princeton University Press.
- Vaughan, D. (1996). *The Challenger launch decision: Risky technology, culture and deviance at NASA*. Chicago, IL: The University of Chicago Press.
- Vicente, K. (1999). *Cognitive work analysis*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Wilkinson, S. (1994). The Oscar November Incident. *Air & Space*, February-March, pp. 80-87.
- Woods, D. D., & Cook, R. I. (2002). Nine steps to move forward from error. *Cognition, Technology and Work*, 4, 137-144.