Taylor & Francis
Taylor & Francis Group

What is rational about killing a patient with an overdose? Enlightenment, continental philosophy and the role of the human subject in system failure

Sidney W.A. Dekker*

Key Centre for Ethics, Law, Justice and Governance, Griffith University, Brisbane, Australia

This paper raises the issue of ergonomics' role in giving primacy to fully rational individual human actors in the creation of system failure, despite its commitment to see action as constrained by design and operational features of work. Reflecting on recent contributions to the journal, ergonomics' dilemma is considered against Enlightenment assumptions about individual human reason as the route to truth and goodness and its critics in continental philosophy. There is a pervasive, unstated pact here. What ergonomics chooses to call certain things (violations, errors, non-compliance, situation awareness) not only simultaneously affirms and denies full rationality on part of the people it studies, it also coincides with what the West sees as scientific, true and instrumental. Thus, ergonomics research legitimates its findings in terms it is expected to explain itself in. But by doing so, it reproduces the very social order it set out to repudiate.

Statement of Relevance: Ergonomics' choice of words (violations, errors, non-compliance) at once affirms and denies full rationality on part of the people it studies, reproducing the very social order it is meant to question and change.

Keywords: Enlightenment; ergonomics; error; continental philosophy; individualism; rationality

In 2009, pharmacist Eric Cropp was convicted for failing to check a chemotherapy solution that had too much sodium chloride. It killed a 2-year old patient through hypernatremia (excess sodium relative to free water in the body). As usual, a variety of factors contributed, including computer problems (ISMP 2009). Yet the human subject, the individual, was accorded a central role in this system failure. The subsequent trial stressed how Cropp could, and should, have prevented the solution from passing to the patient.

A consistent commitment of ergonomics has been to understand the system surrounding the production of error and expertise (Hollnagel *et al.* 1999, Woods and Cook 2002). The 'systems approach' is now a common term both in the field and its application areas (Reason 1995, Waterson 2009, Goh *et al.* 2010, Leveson 2011). Despite this, individual agency at the sharp end remains a popular site for accountability and intervention (Cook and Nemeth 2010). Recently in this journal, Waterson conducted a review of the patient safety literature that confirmed this as a trend (Waterson 2009). Although some of the studies actively looked at teams (17%) or organisations (14%), 98 of the 360 articles reviewed addressed the individual level of analysis, focusing, for example, on

human error. Similarly, medical adverse event investigations are known to stress individual agency and responsibility (Sharpe 2004, Berlinger 2005, Cook *et al.* 2008). Even fields of practice that have been open to ergonomics from the beginning, such as aviation (Roscoe 1997), have difficulty escaping this focus. Between 1999 and 2006, 96% of US aviation accidents were attributed in large part to the flight crew. In 81%, people were the sole reported cause (Holden 2009). Accident probes often conduct analyses of people's decision making as if it were driven by rational, fully informed choices, concluding that they either must have been amoral calculators, who prioritised production or personal goals over safety (Vaughan 1999), or slackers, who somehow failed to invest in their own full rationality, leading to a reduced 'situation awareness' (Parasuraman *et al.* 2008) or shortcuts that are called 'violations' (Reason 1990).

The contrast between such images of full rationality—which people in principle can achieve and in hindsight should have achieved—stands in sharp contrast with 70 years of ergonomics work. Ergonomics stresses the influence of context on human action. Pioneering work by Fitts, Chapanis, Senders and others demonstrated how the ideal of fully rational, regulative human action is constrained by

---

*Email: s.dekker@griffith.edu.au

design and operational features (Fitts and Jones 1947, Chapanis 1970, Senders and Moray 1991, Roscoe 1997). Design things in certain ways and some errors become almost inevitable. It became the basis for the ergonomic enterprise: human error is systematically connected to features of people's tools and tasks. Simon's critique of rational decision making lent theoretical weight (Simon 1959). Using a mix of computational argument (the potential problem space is far larger than the cognitive capacity of the decision maker), empirical observations and thought experiments (Simon 1957, 1959, 1969), he argued that human rationality is not omniscient but 'bounded'. The assumptions of behavioural theory about human capabilities, he proposed, should be far weaker than those of rational choice theory and make modest, realistic demands on the knowledge and computational abilities of human agents (Newell and Simon 1972).

Discoveries about hindsight and other decision biases (Fischhoff 1975), cognitive fixation (de Keyser and Woods 1990) and social dynamics of decision formation (Janis 1982) added questions about the accuracy and relevance of rational choice assumptions in human work (Klein 1989, Orasanu and Connolly 1993). 'Bounded' rationality became the new ortho-doxy until it became clear that it both invoked and denied full rationality (since rationality can be 'bounded' only in comparison to some full ideal). It was gradually replaced by the label 'local' rationality (Woods *et al.* 1994). The notion of local rationality fit characterisations of cybernetics (Ashby 1956), an important tributary to ergonomics. Knowledge is intrinsically local and subjective, an imperfect tool used by an agent to help it achieve its goals. The agent not only does not need perfect rationality, it can never achieve it. It locally senses its inputs and notes its outputs (actions), inducing regularities that hold in its environment (Heylighen *et al.* 2006). What these agents do is locally rational—assessments and actions make sense given the agent's active goals, current knowledge and focus of attention. This premise of local rationality forms the basis for much ergonomics research (Amer-ican Medical Association 1998). Ergonomics thus de-centred the subject, constructing it as a local player influenced by system- and organisational factors that constrain what that subject might see as rational at any time (Vaughan 1996). For ergonomics, the subject was a product rather than the creator of the social order.

This raises a critical question. If 70 years of compelling theory and empirical progress cannot be persuasive with regard to the limits of human rationality in complex, dynamic settings, then what can be? What is ergonomics doing wrong?

Of course, there is quite a lot to be overcome. The correspondence bias in attribution theory makes people draw inferences about a person's unique and enduring dispositions from behaviours that can be entirely explained by the situations in which they occur (Gilbert and Malone 1995). It remains strong and is consistent with individualism and Western notions of agency and accountability (Choi *et al.* 1999, Bauman and Skitka 2010). Just as there are well-documented psychological biases, larger societal shifts surely account for some of the regression away from local rationality as an explanation for people's role in system failure. Growing risk aversion and declining familiarity with risk in the West seem coupled to the notion that system breakdowns no longer are meaningless coincidences of space-time, but represent failures of risk management (Green 2003, Dekker 2011a). Criminalisation of human error is on the rise in worlds ranging from aviation to healthcare, affirming the central role of the human subject as rational decision maker who can be held accountable for having failed to manage risk appropriately (Dekker 2011a). The *British Medical Journal* has even banned the use of the word 'accident' in 2001 for describing adverse medical events—implying that each such event has (potentially preventable) human agency behind it and therefore represents nothing 'accidental' (Davis and Pless 2001). Rational choice assumptions thus retain preferential status in theories of organisational decision making (Page 2007). It puts practical and political limits on ergonomics to make palpable progress (Mason 2010). All this may—through societal proclivity, political impotence, editorial fancy, modernist hubris—reaffirm the central role of the subject. It is the individual subject who has the potential to attain full rationality and make decisions to not err and break anything. As the mother of the 2-year old patient said in court while turning to Eric Cropp: 'You were the only person who could have prevented this death from happening, and you didn't do it. You killed my baby' (ISMP 2009).

These are considerable obstacles. They might be overcome by doing more of what ergonomics is already doing, or by trying even harder. But perhaps subtly and pervasively, the possibility of full rationality, and the subject's central role in attaining it, gets reproduced and reconfirmed in the language of ergonomics itself. Ergonomics research typically explains itself in terms of a Western regulative moral ideal with its emphasis on individualism, volition and universal rationality (Althusser 1984). These explanations also tend to coincide with what the West sees as scientific (Wilkin 2009). Testimony to this is the asserted usefulness of a concept such as 'situation awareness' in ergonomics (Parasuraman *et al.* 2008). Whatever limited awareness was present inside an operator's mind, it can always be contrasted against a sort of 'ground truth' (ibid, p. 144) or fully rational

ideal that the operator could or should have achieved. This represents a naïve Newtonian scientism: total knowledge of the world is achievable; the world is 'out there' as an object entirely separable from observers. People will know the truth if they are fully rational, once the correspondence between the picture in their mind and reality in the world is perfect (Dekker and Lützhöft 2004).

Ergonomics has not shied away from a language that similarly resorts to people's potential full rationality and that appeals to individual action, agency and responsibility to achieve it. Terms such as 'error' and 'violation' have been popularised (Reason 1990) and remain useful handmaidens in a variety of organisational, political and psychological processes (Cook and Nemeth 2010). The snare is that such terms at once deny and invoke full rationality. Nyssen's and Cote's (2010) ergonomics studies in pharmacy, for example, show that the rationality of operators is constrained by cognitive and social influences—a finding consistent with ergonomic orthodoxy. Yet the results are counted and expressed in terms of 'non-compliance' and 'violations': the gap between procedure and actual work can be closed by taking into account people's motivational factors (Nyssen and Cote 2010). These are rhetorical commitments that belong to a normative universe, where individuals are non-compliant relative to something that represents full compliance, something that is fully rational. And what is more fully rational than a procedure that is the carefully considered, rationally designed, best thought-out way to do a job? Such thinking goes back to the pre-historical project of scientific management (Wright and McCarthy 2003): full rationality and perfect correspondence between action and rule are in principle achievable.

In the dialectic between universal rationality of the subject on the one hand and a constrained, merely local bit-player on the other, ergonomics is in good company. The former invokes Enlightenment ideals about universal rationality and the centrality of the subject. The Enlightenment saw people as self-contained individuals, driven by the power of reason—a position that has become synonymous with scientific, rational thought in the West. In Descartes' *Meditations* (1641), human rationality became seen as the basis for systematic, true knowledge about the world. Individual human capacities and concerns started taking precedence, emphasising the centrality of the subject and its rational ways of solving problems. Even the Reformation valorised individual mores and deeds, away from submission to Church tradition. Rights, privileges and duties were increasingly ascribed to the subject, the individual, whose self-realisation, morality and accountability have become central

preoccupations of Western thought and action. With their full rationality, humans could do the right thing.

Questions about this image of fully rational, centrally positioned subjects were not far behind, however. They predated ergonomics' own construction of the subject as de-centred and merely locally rational. The intellectual tradition known as continental philosophy has taken Enlightenment assumptions about full rationality and its central subject to task ever since Kant (1724–1804). Hegel (1770–1831), for instance, emphasised history and social context as hugely influencing the subject's rationality. Similarly, Marx (1813–1883) saw the human subject not as a free, universally rational being but rather as an ensemble of social relations. What any person can see, or understand, or construe as rational, is dependent on where and how they are configured in the larger web of society (West 1991). Freud (1856–1939) suggested how human rationality is constrained and influenced by the unconscious mind, which allows only a distorted, partial human consciousness, ever subjected to the drives and repressive vetoes of the id (Freud 1950). Foucault (1926–1984) demonstrated that the human subject is a finite and historical entity; that what the subject can know is always the relative and questionable expression of a particular constellation of relationships in which that subject is configured (Foucault 1970). Continental philosophy thus put the human subject beyond the reach of pure reason, or full rationality. Instead, the subject is historically and locally configured, with a rationality that is constrained by social position and mental workings. This stands in contrast to the universalising and individualist impulses of Enlightenment and modernism, which to this day reach deeply into popular Western notions of science and truth (Wallerstein 1996).

Ergonomics is not immune to such influences. It too might need to produce research that is objectifying (e.g. counting instances of violations or errors away from the subjective context in which these actions occurred), thereby supposedly generating timeless, universal and instrumental knowledge. The language that ergonomics helped give to the death of Eric Cropp's little girl (errors, violations, non-compliance) does not just reflect an empirical world, if it does that at all. It shapes the world everyone can know, in ways that both facilitate and constrain action (Healy 2003). If a medication death was preceded by non-compliance, then it was preventable. If it was preventable, the human subject might have mustered the rationality to prevent it. Ergonomics supplies linguistic resources that not only reproduce Enlightenment notions of individualism, rationality, morality and accountability, but that also delimit the repertoire of countermeasures and leave the traditional targets of ergonomic system intervention

under-explored. Based on the findings of his review, Waterson argued that research really needs to move away from a concentration on errors and towards an examination of the connections between systems levels (Waterson 2009). In a Western world where assumptions about the centrality of the subject and full rationality are still coincident with science, truth and instrumentality, they easily get retained in ergonomic projects that set out to repudiate them. It reproduces an existing order, where, for instance, rational choices about procedure-following can guarantee safety (Dekker 2003). Such a reproduction might remain largely transparent to the ergonomic community and closed to critique precisely because its discourse is so ordinary, self-evident and commonsensical.

Fortunately, the last 20 years in ergonomics theory have seen a gradual move away from the use of terms that denote individual rationality and agency. From arguments for a system approach that still radiate outward from errors and violations as their core (Reason 1995), a more wholesale shift is evident in recent work that promotes new vocabularies and eschews individual attributions of success and failure. This includes the movement of resilience engineering (Hollnagel et al. 2006), which models how systems adapt to their changing environments and absorb challenges to their integrity. And there are exciting applications of complexity theory that see system failure as growing out of relationships, not failed components (Dekker 2011b, Leveson 2011). Also encouraging is the emergence, if ever so slight, of what could be called a critical ergonomics that examines its mainstream positivist and empiricist philosophical position (Dekker et al. 2010). Actively theorising and critiquing ergonomics as a mix of politics, technology and psychology, it suggests that ergonomics' reproduction of Newtonian science not only undermines its ability to properly theorise the social world (Flach et al. 2008), but turns it into an ideological instrument whose social use largely reflects the interests of dominant social institutions (Wilkin 2009). Challenging assumptions through a critical ergonomics, and changing languages to construct a different empirical and ontological field of inquiry, are necessary developments. They might help unseat those hegemonic assumptions about rationality and individual agency that have so far safeguarded the supremacy of the subject in Western ergonomics.

References

Althusser, L., 1984. *Essays on ideology*. London: Verso.

American Medical Association, 1998. *A tale of two stories: Constrasting views of patient safety*. Chicago, IL: American Medical Association.

Ashby, W.R., 1956. *An introduction to cybernetics*. New York: J. Wiley.

Bauman, C.W. and Skitka, L.J., 2010. Making attributions for behaviors: the prevalence of correspondence bias in the general population. *Basic and Applied Social Psychology*, 32, 269–277.

Berlinger, N., 2005. *After harm: medical error and the ethics of forgiveness*. Baltimore: Johns Hopkins University Press.

Chapanis, A., 1970. Human factors in systems engineering. *In*: K.B. de Green, ed. *Systems psychology*. New York: McGraw-Hill, 51–78.

Choi, L., Nisbett, R.E., and Noenzayan, A., 1999. Causal attribution across cultures: variation and universality. *Psychological Bulletin*, 125, 47–63.

Cook, R.I. and Nemeth, C.P., 2010. 'Those found responsible have been sacked': some observations on the usefulness of error. *Cognition, Technology and Work*, 12, 87–93.

Cook, R.I., Nemeth, C., and Dekker, S.W.A., 2008. What went wrong at the Beatson Oncology Centre?. *In*: E. Hollnagel, C.P. Nemeth, and S.W.A. Dekker, eds. *Resilience engineering perspectives: Remaining sensitive to the possibility of failure*. Aldershot, UK: Ashgate Publishing Co.

Davis, R.M. and Pless, B., 2001. BMJ bans 'accidents'. *British Medical Journal*, 322, 1320.

de Keyser, V. and Woods, D.D., 1990. Fixation errors: failures to revise situation assessment in dynamic and risky systems. *In*: A.G. Colombo and A. Saiz de Bustamante, eds. *System reliability assessment*. The Netherlands: Kluwer Academic, 231–251.

Dekker, S.W.A., 2003. Failure to adapt or adaptations that fail: contrasting models on procedures and safety. *Applied Ergonomics*, 34 (3), 233–238.

Dekker, S.W.A., 2011a. The criminalization of human error in aviation and healthcare: a review. *Safety Science*, 49, 121–127.

Dekker, S.W.A., 2011b. *Drift into failure: From hunting broken components to understanding complex systems*. Farnham, UK: Ashgate Publishing Co.

Dekker, S.W.A. and Lützhöft, M., 2004. Correspondence, cognition and sensemaking: a radical empiricist approach to situation awareness. *In*: S. Banbury and S. Tremblay, eds. *Situation awareness: A cognitive approach*. Aldershot, UK: Ashgate Publishing Co.

Dekker, S.W.A., et al., 2010. Epistemological self-confidence in human factors research. *Journal of Cognitive Engineering and Decision Making*, 4 (1), 27–38.

Fischhoff, B., 1975. Hindsight ≠ foresight: the effect of outcome knowledge on judgment under uncertainty. *Journal of Experimental Psychology: Human Perception and Performance*, 1 (3), 288–299.

Fitts, P.M. and Jones, R.E., 1947. *Analysis of factors contributing to 460 'pilot error' experiences in operating aircraft controls.No. TSEAA-694–12*. Dayton, OH: Aero Medical Laboratory, Air Material Command, Wright-Patterson Air Force Base.

Flach, J.M., Dekker, S.W.A., and Stappers, P.J., 2008. Playing twenty questions with nature: reflections on the dynamics of experience. *Theoretical Issues in Ergonomics Science*, 9 (2), 125–155.

Foucault, M., 1970. *The order of things: An archaeology of the human sciences*. London: Tavistock Publications.

Freud, S., 1950. Project for a scientific psychology. *The standard edition of the complete psychological works of Sigmund Freud*, vol. 1. London: Hogarth Press.

Gilbert, D.T. and Malone, P.S., 1995. The correspondence bias. *Psychological Bulletin*, 117 (1), 21–38.

Goh, Y.M., Brown, H., and Spickett, J., 2010. Applying systems thinking concepts in the analysis of major incidents and safety culture. *Safety Science*, 48 (3), 302–309.

Green, J., 2003. The ultimate challenge for risk technologies: controlling the accidental. *In*: J. Summerton and B. Berner, eds. *Constructing risk and safety in technological practice*. London: Routledge.

Healy, S., 2003. Epistemological pluralism and the 'politics of choice'. *Futures*, 35, 689–701.

Heylighen, F., Cilliers, P., and Gershenson, C., 2006. *Complexity and philosophy*. Brussel, BE: Vrije Universiteit Brussel.

Holden, R.J., 2009. People or systems: to blame is human. The fix is to engineer. *Professional Safety*, 12, 34–41.

Hollnagel, E., Kaarstad, M., and Lee, H., 1999. Error mode prediction. *Ergonomics*, 42 (11), 1457–1471.

Hollnagel, E., Woods, D.D., and Leveson, N.G., 2006. *Resilience engineering: Concepts and precepts*. Aldershot, UK: Ashgate Publishing Co.

ISMP, 2009. August 27. Ohio government plays whack-a-mole with pharmacist. *ISMP Medication Safety Alert*, 14, 1–3.

Janis, I.L., 1982. *Groupthink*, second edition. Chicago, IL: Houghton Mifflin.

Klein, G.A., 1989. Do decision biases explain too much? *Human Factors Society Bulletin*, 32 (5), 1–3.

Leveson, N.G., 2011. Applying systems thinking to analyze and learn from accidents. *Safety Science*, 49 (1), 55–64.

Mason, S., 2010. Understanding human error in mine safety. *Ergonomics*, 53 (12), 1505–1506.

Newell, A. and Simon, H.A., 1972. *Human problem solving*. Englewood Cliffs, NJ: Prentice-Hall.

Nyssen, A. and Cote, V., 2010. Motivational mechanisms at the origin of control task violations: an analytical case study in the pharmaceutical industry. *Ergonomics*, 53 (9), 1076–1084.

Orasanu, J.M. and Connolly, T., 1993. The reinvention of decision making. *In*: G.A. Klein, J.M. Orasanu, R. Calderwood, and C.E. Zsambok, eds. *Decision making in action: Models and methods*. Norwood, NJ: Ablex, 3–20.

Page, S.E., 2007. Making the difference: applying a logic of diversity. *Academy of Management Perspectives*, 11, 6–20.

Parasuraman, R., Sheridan, T.B., and Wickens, C.D., 2008. Situation awareness, mental workload and trust in automation: viable, empirically supported cognitive engineering constructs. *Journal of Cognitive Engineering and Decision Making*, 2 (2), 140–160.

Reason, J.T., 1990. *Human error*. New York: Cambridge University Press.

Reason, J.T., 1995. A systems approach to organisational error. *Ergonomics*, 38 (8), 1708–1721.

Roscoe, S.N., 1997. The adolescence of engineering psychology. *In*: S.M. Casey, ed. *Human factors history monograph series, volume*, vol. 1. Santa Monica, CA: Human Factors and Ergonomics Society.

Senders, J.W. and Moray, N., 1991. *Human error: Cause, prediction, and reduction*. Hillsdale, NJ: L. Erlbaum Associates.

Sharpe, V.A., 2004. *Accountability: Patient safety and policy reform*. Washington, DC: Georgetown University Press.

Simon, H.A., 1957. *Models of man: social and rational; mathematical essays on rational human behavior in a social setting*. New York: Wiley.

Simon, H.A., 1959. Theories of decision-making in economics and behavioral science. *The American Economic Review*, 49 (3), 253–283.

Simon, H.A., 1969. *The sciences of the artificial*. Cambridge: MIT Press.

Vaughan, D., 1996. *The Challenger launch decision: Risky technology, culture, and deviance at NASA*. Chicago: University of Chicago Press.

Vaughan, D., 1999. The dark side of organizations: mistake, misconduct, and disaster. *Annual Review of Sociology*, 25, 271–305.

Wallerstein, I., 1996. *Open the social sciences: Report of the Gulbenkian Commission on the Restructuring of the Social Sciences*. Stanford, CA: Stanford University Press.

Waterson, P., 2009. A critical review of the systems approach within patient safety research. *Ergonomics*, 52 (10), 1185–1195.

West, D., 1991. *An introduction to continental philosophy*. London, UK: Polity Press.

Wilkin, P., 2009. The ideology of ergonomics. *Theoretical Issues in Ergonomics Science*, 11 (3), 230–244.

Woods, D.D. and Cook, R.I., 2002. Nine steps to move forward from error. *Cognition, Technology & Work*, 4 (2), 137–144.

Woods, D.D., et al., 1994. *Behind human error: Cognitive systems, computers and hindsight*. Dayton, OH: CSERIAC.

Wright, P.C. and McCarthy, J., 2003. Analysis of procedure following as concerned work. *In*: E. Hollnagel, ed. *Handbook of cognitive task design*. Mahwah, NJ: Lawrence Erlbaum Associates, 679–700.