This is the Author's Original version. The published version contains corrections to errors and can be accessed at: https://www.cambridge.org/core/journals/philosophy/article/ethics-within-engineering-by-wade-l-robison-bloomsbury-publishing-2017-pp-xvi-281-2399-isbn-9781474286053/4638F5C37834ECBBDD137D5353F56A3F

Ethics Within Engineering advances Wade Robison's weighty, interdisciplinary goal of "chang[ing] the way in which ethics is taught in engineering" (xvi). Robison's 'way' proceeds along two dimensions. Firstly, he aims to relocate ethical discussions from "add-on" courses (xvi) and the endings of engineering syllabi to "the first chapter of any engineering text . . ." (252). Secondly, and more controversially, Robison targets the philosophers who teach ethics to engineers; he recommends a pedagogy free of formal theories of virtue, duty, or the utilitarian calculus (259-260).

Chapter 1 exposes the universal, human frustration with "designs that provoke errors" (2), such as toasters that break if the user lifts the handle to eject the bread instead of dialing off the heat. Such errors of wasted time, bread, money, and the very *risk* of waste Robison calls "harms" (3-4), and he claims: "*It is a basic moral principle that we should cause no unnecessary harm* . . ." (1). One species of unnecessariness, on Robison's account, is being foreseeably preventable. He surveys foreseeable harms anticipated or not in the designs of additional, unglamorous artifacts like toothpicks and doorknobs.

With the concept of error-provocation thus established, chapter 2 scrutinizes accusations of operator error. "[I]n any accident," Robison claims, "there are two other variables . . . besides the operator" (29, italics removed). Those variables are (a) the harmful "circumstances" that may prove impossible to reasonably foresee (e.g. "black ice"), and (b) the artifact whose design is "error-provocative" or not (29, italics removed). Robison demarcates (a) from (b) to prepare his analysis in chapter 3 of the mostly fatal crash of the 163-passenger American Airlines Flight 965 in Colombia in 1996. That crash resulted from a software incongruity, in which the pilot or co-pilot's entering the letter "R" into the navigation computer automatically set the aircraft's destination to Bogota (44). Most other single-letter depressions of the navigation computer's keyboard generated a list of beacons sorted by distance, whose navigation code began with the letter pressed (44). Bogota's beacon was named "Romero," and the pilots' target, Cali, was named "Rozo" (44-45). Upon receiving the letter "R," the navigation computer selected

Romero, which was “100 miles” from the nearby Cali (45), and crashed the airplane into a mountain, killing 159 people (46, 44). Robison points out that the *New York Times* and “[t]he airline” blamed the crash on “the pilot’s error” (46). But Robison objects that the navigation software was error-provocative. The special behavior of the computer with respect to an R-depression “creates a problem that we can predict,” Robison claims, because “we are likely to forget” spurious exceptions to rules (49). Robison does not exonerate the pilot “completely” (53), but does attribute “moral failure” to the software engineers who allowed such an inconsistency into production (55).

Chapter 4 returns to “mundane” design problems like the arrangement of stove burners and knobs (57, italics removed). Robison claims that the “first decision” executed in designing a stove-top is the decision to make the stove “eas[y]” to use (58). This value of ease imparts a moral “*ought*” to the community of stove-users, Robison argues, because valuing ease amounts to valuing the avoidance of harm (59). Thus Robison denies that ethics is added-into the design project, as if by pious reflection; to the contrary, he insists that “conditional claims” about the operational parameters of an artifact may themselves entail “normative judgment[s]” (59).

Robison reinforces the integrity of ethics to engineering in chapter 5, where he relates agential intention to moral blame. He recognizes a “gold standard” for ascribing blame when agents possess sound minds, know the harmfulness of an act, and nevertheless deliberately execute it (99). Via some alarming examples, however, Robison demonstrates that we also blame agents who possess no bad intent, such as the incompetent (surgeons who cannot identify organs) and the unprofessional (engineers who could double-check their calculations, but do not). Hence Robison charges the software engineers behind the aforementioned Flight 965 crash with incompetence for not foreseeing the computer program’s error-provocativeness (115), and with unprofessionalism for not eliminating the fault during testing (117). Ultimately Robison identifies five moral relations into which engineers enter (206-207), and his discussion of intention names three of them: (i) “Role morality,” or a competence at applying mathematics and physics to create functional artifacts; (ii) a norm “Internal” to engineering that design solutions should minimize harm; and (iii) “External” responsibilities that engineers have not to harm their employers and clients (118-119).

Chapter 6 introduces the fourth of Robison’s moral relations for engineers, (iv) the “Aspirational” attitude of continuous improvement, or the spurning of mediocrity with respect to design solutions (138). He implies, however, that the best aspirations remain inert if engineers "fail to understand a design problem fully . . ." (139). 'Fully' understanding, on Robison's account, includes anticipating extra-laboratory conditions to which users will subject the artifact. One infamous model of X-ray machine, for example, deenergizes into a position designed to protect the machine from cleaning staff who need to maneuver around it, but such a position also happens to crush any human patient who has not exited the scanning table. Robison surmises that had the shutdown software been written by engineers cognizant of the fact that the X-ray operator sometimes lacks line-of-sight to the patient, an occupancy sensor might have been added to the scanning table (144).

Robison acknowledges in chapter 7, on the other hand, that not all harms arise from users provoked into making *mistakes*. The Ford cruise control switch seal fatigued unexpectedly over time due to normal operation of the brakes, and spontaneous fires resulted (157). By my reading, Robison implies that 'aspirational', continuous improvement means anticipating such anomalous, system interactions. He similarly admonishes engineers to "look upstream" to reduce harms from material selection, disposal, and transport (183).

The penultimate Chapter 8 identifies necessary but insufficient (195) conditions for professionalism: knowledge, skills (195), certification (196), and a "special set of moral relations . . ." (197). Such moral relations Robison parses as responsibilities to perform various species of action (202). Physicians probe physically, for example, whereas attorneys probe only dialectically (197-198). Robison recognizes that ethical conflicts may arise between an agent's role *qua* "professional" and her role *qua* "employee," but he absconds from analyzing such conflicts (205). He instead groups the distinctly professional, act-specific relations as the fifth kind that engineers contract, i.e. (v) "Social" relations that fall outside the purview of relations (i) through (iv) (206). Chapter 9 reiterates Robison's thesis.

Despite *Ethics Within Engineering*'s prescience in addressing serious and ubiquitous, real-world harms like the 1992 Subaru shoulder harness (171) and Baxter's inconspicuous labeling of adult and infant blood thinners (233 ff.), I hesitate to recommend the book. Granted, I think that Robison achieves his goal; by the final page he has amassed a mountain of examples and analogies supporting the claim that engineering is value-laden, and I have no inclination to disagree with him. What I question is (1) why I (*qua* hypothetical, aspiring engineer) should care about the five types of moral relations that Robison forecasts, and (2) the effectiveness of Robison's pedagogy.

Per (2), some of Robison's adjurations come across as naive. He claims, for example, that "[e]ngineers . . . need at the least to learn what would count as the right sort of answer to a calculation, one in the ballpark of answers" (218). As an engineer-turned-philosophy-instructor myself, I can attest that no undergraduate qualified to *declare* a major in engineering has passed the prerequisite courses in calculus and physics without years of constant reliance on ballpark-checking. Just as *Ethics Within Engineering* tells undergraduates what they might not need to hear, so I question the long-term impact of Robison's somewhat sermonizing style.

Hence problem (1). Why be moral? Robison tacitly appeals throughout the book to precedents of virtuous agents and actions (121, 136), even quoting Aristotle a number of times (106, 200). In the end, however, I only reckon that I should be moral because the law will get me, or because Robison's opinion resonates so strongly after 260 pages. Required reading of *Ethics Within Engineering* may condition pre-professional novices to an important and unfamiliar perspective, but not more so, I imagine, than other college pep-talks. By contrast, human resource departments train workers to identify 'the four personality types', rather than hope that a charismatic speaker will hypnotize the workers into getting along. Why not challenge the highly analytic science majors to identify the ethical paradigms to which people frequently resort, and reveal to those undergraduates some of the limitations of each paradigm? Robison offers only the briefest replies to this question, claiming that "if philosophers cannot agree [on which ethical paradigm wins], we can hardly expect engineers to digest these theories and make a rational and moral choice between them . . ." (260). I reply that 'digestion' is a straw man. In contrast, I find even the *attempt* at sophisticated justification of moral decisions to be itself a kind of action-brake that retards agents long enough to think twice — as do the safety-slogan door mats at factory entrances. My method might not motivate moral action any better than Robison's, but I can appeal to engineers' analytical proclivities in hopes of slowing the engineers down. I do not expect them to walk around thinking about ethical-theoretic minutia all day; indeed, engineers should not, they should be thinking about artifacts designed according to safety-conscious processes. But if the engineers are challenged to understand that one more analysis can *always* be performed, i.e. the *ethical* analysis, then the action-brake (thanks to philosophical training) will be doing its job. Because of its disowning of formal theories, *Ethics*

Within Engineering does not quite comport with my teaching strategy, even if the book is strongly accessible with a rich bibliography of recent incidents.