The Ethics of Acceptable Safety

Ibrahim Habli[1], Tim Kelly[1], Kevin Macnish[2], Christopher Megone[2], Mark Nicholson[1], Andrew Rae[3]

[1] Department of Computer Science, University of York, UK

[2] Inter-Disciplinary Ethics Applied Centre, University of Leeds, UK

[3] Safety Science Innovation Lab, Griffith University, Australia

Abstract *Engineers of safety-critical systems have a duty to address ethical issues that may arise in the development, assessment, operation and maintenance of these systems. Dealing with ethical dilemmas during safety risk assessment is particularly challenging, especially when making and justifying decisions concerning risk acceptability. This is particularly complicated by organisational issues and contractual limits that do not necessarily align with the boundaries of ethical responsibility. In this paper, we explore some of these dilemmas and discuss the duties of engineers to identify, analyse and respond effectively to ethical concerns about safety risk decisions. We illustrate these through short case studies that highlight particular issues relating to the ethics of safety advice, safety and cost tradeoffs, novel technologies and institutional support.*

1 Introduction

Any discussion of engineering ethics risks belabouring the obvious and ignoring the true challenges. No engineer (we hope) sets out to be evil. All have, at some stage of their career, agreed to follow a carefully worded code of conduct. Why do we, as professional educators of engineers, feel that it is necessary to write a paper about the ethics of safety?

The engineering of safety-critical systems is a constant process of ethical decision making. Some of these decisions are subtle, such as balancing cost, performance and reliability in the selection of a component, or choosing the right wording for a customer memo. Other decisions are explicit and life-changing, such as deciding whether to persevere in a dysfunctional organisation in the hope of eventually making a positive difference.

Throughout this paper we aim to provide some insight into what makes safety ethics difficult. We first set out principles on which ethical decision making may be founded by introducing the Royal Academy of Engineering's Statement of Eth-

ical Principles. We then show how these principles are stretched, challenged and sometimes directly threatened by contractual obligations, organisational capability, competing ethical concerns, and the uncertain nature of risk acceptability. We explore and illustrate these complications through a series of short case studies. Finally, we conclude with some practical suggestions for ways forward.

This paper does not provide simple answers. As we highlight in some of the examples, we are fellow-travellers on the safety journey, not beacons or signposts. Safety ethics is just as problematic when educating and advice-giving as it is in the design, maintenance and operation of safety-critical systems. We hope that by explicitly acknowledging the challenges and complications our work will help engineers reflect on their own practice and be more confident in recognising and resolving ethical problems.

2 Royal Academy of Engineering's Statement of Ethical Principles

An engineer applies scientific and technical knowledge to address problems within their domain of expertise. As members of a professional discipline, they are required to take account of ethical considerations in their work. What exactly does it mean to 'take account of' ethical considerations, and what can be done to enhance and develop engineering practice in this regard?

In the UK, various engineering institutions have developed slightly differing codes of conduct or codes of ethics. In 2005, the Royal Academy of Engineering, working in conjunction with the Engineering Council and a number of leading engineering institutions, developed a Statement of Ethical Principles (RAEng 2011a). This was partly to address issues of uniformity, but more importantly to raise the profile of ethical considerations amongst engineers. The aim was to have a succinct statement to which all practising engineers could subscribe. The Statement presented a set of four principles designed to cover all of the fundamental ethical considerations that are most prominent in an engineer's day to day work[1]. The principles are:

- Accuracy and Rigour;
- Honesty and Integrity;
- Respect for Life, Law and the Public Good; and,
- Responsible Leadership, Listening and Informing.

Engineers are a practical group of people. Faced with such brief and rather abstract principles, they are going to ask how exactly these should be used to guide behaviour. In the short launch document for the Statement there is a brief elabora-

[1] As noted many professional engineering institutions also have codes of ethics, but this statement has been endorsed by both the Engineering Council and leading institutions, so constitutes a good starting point for those in all areas of safety engineering.

tion of each of the principles with several short bullet points. Even this elaboration is still quite abstract, and so the question remains as to how exactly these general principles can help an engineer meet high ethical standards.

One of the factors that makes it difficult to apply the principles is that it will often be the case in real-life situations that the principles intersect. It will not simply be a case of recognising one principle and working out how to address its requirements in the particular circumstance faced, but of considering what to do given that more than one principle is in play. Furthermore the principles may sometimes conflict, making a decision even more difficult.

Ethicists will often talk of weighing or balancing different considerations in order to arrive at good judgements. For those in technical disciplines it is easy to assume that this will involve some kind of algorithm or a precise mathematical process. However, as Aristotle pointed out a long time ago, although there is truth in ethics as there is in mathematics, the kind of precision that is expected in mathematics is not appropriate in ethics. Thus, although it is tempting to seek a grid or a formula which reduces the complexity of ethical choice, the apparent precision achieved is gained at the cost of over-simplifying the decision process. An algorithmic approach to ethics leads to significant factors being ignored or not properly taken account of.

So, the intersection of ethical principles, including possible conflict between them, is one challenge to their ready use in decision making. Another is the fact that the principles are not self-interpreting rules. Reflection is required in order to work out what any given principle might mean in context. Some ethical principles may sometimes in practice require absolute prohibitions on certain types of behaviour, whilst on other occasions they may point to ideals of behaviour to which we might aspire but realistically must inevitably fall short of.

Consider the first principle which highlights the need for accuracy and rigour. How much accuracy and rigour is it reasonable to expect on any given occasion? Accuracy and rigour are 'limit' concepts; they admit of ideals to which our ordinary practice can approach, but never fully instantiate. It will always be open to question how much time, effort and cost should be expended on achieving ever greater accuracy. A sub-principle here advises that engineers "should identify and evaluate and, where possible, quantify risk" (RAEng 2011) but the identification of risks can be an endless task if one includes every single risk which an engineering intervention might introduce. We touch on this point further below.

A final issue to mention here is that the Royal Academy's ethical principles must be applied in context. This means they will need to be acted on at times when the practitioner will be facing commercial considerations of cost and contract, personal issues such as their own needs and the demands of family and friends, and matters of organisational culture. In applying the principles all these factors complicate matters. We take these ideas further in the next section, but one thing to note immediately is that, given the importance of organisational culture, effective implementation of the principles requires that the organisations in which engineers work find suitable ways to embed these principles within that organisational culture.

Despite these difficulties, the Statement of Ethical Principles provides a solid starting point for thinking about the ethics of safety. The genuine ethical concern that we feel when the principles cannot be simultaneously met, or when their ideals cannot be instantiated, indicates that they speak to what it means to practise engineering ethically. Recognising the principles, and identifying when they are challenged by situations and actions is the first step towards ethical behaviour.

3 Day to Day Application of Principles for Safety Management

Having introduced both the four Ethical Principles and some of the challenges that accompany them, in this section we expand upon these challenges by applying these principles to safety management within real world organisations. These challenges include recognising the fact that one is presented with an ethical decision; deciding what is reasonable; dealing with uncertainty; reconciling competing duties; making money; everyday considerations; and choosing the greater good.

Recognising the Ethical Nature of Decisions
Engineers are seldom presented with a stark choice between ethical and unethical options. Problems involving pulling a lever to kill one person instead of five people appear in philosophy tracts, not requirements documents (Thomson 1976). Often even the fact that a choice is being made is not evident. When an engineer records the appropriate Safety Integrity Level (SIL) or Design Assurance Level (DAL) for a component, this may appear to be a deterministic calculation. In fact since SILs and DALs determine assurance effort, this is a cost/safety tradeoff. As the engineer makes assumptions and chooses parameters under uncertainty, they are exercising the "Respect for Life, Law and the Public Good" principal. As they decide whether it is worthwhile pursuing more information before selecting these values, they are exercising the "Accuracy and Rigour" principle.

Framing engineering or management decisions as ethical choices is often a rhetorical device used to justify a more conservative or expensive option. At other times the decisions are seen as "simply" engineering judgements or management choices, with no explicit ethical content. For example, setting a discount rate or internal rate of return (IRR) is a company strategic decision for making project and investment choices. The IRR can have a significant impact on which mitigations appear feasible or impracticable in a trade-off calculation, as it changes the present value of future costs. Applying the IRR to safety improvements can also result in inter-generational inequity, by valuing future victims less than persons currently alive.

Choosing suppliers is another decision fraught with safety implications. System safety may depend on the reliability of components where there is no direct visibility of their manufacture or quality assurance processes. Selecting a cheap supplier may be a direct cost/safety tradeoff, although it is usually expressed instead as a straight "value-for-money" proposition.

Deciding what is Reasonable

As we foreshadowed when we introduced the four principles, no ethical position is absolute. Yes, public safety is an important value, but that doesn't mean that all enterprise presenting any risk should be abandoned. Instead we look to make reasonableness trade-offs. A product is considered "safe" if the resultant risk is considered acceptable (Lowrance 1976). In the UK "As Low as Reasonably Practicable" (ALARP) recognises that public's right to safety is limited by the mitigations that businesses can be reasonably expected to put in place (HSE 2001). Arguably, there is an ethical *obligation* to treat risks proportionately, to avoid over-emphasis on some hazards at the expense of others presenting greater risk.

Words such as "tolerable", "acceptable", "reasonable" and "practicable" are common currency in safety discussions. Ethical behaviour is not about stepping over a line, but drawing the line appropriately (Dekker 2009). No one would suggest presenting the public with intolerable risk, but we do make frequent decisions about which risks are tolerable and which intolerable.

Reasonableness does not just apply to direct safety, but to related issues such as risk communication. An obligation to share information about risk normally does not extend to revealing commercial secrets. What if a company has invented a safer way of performing an activity? Is it right to treat that as a commercial advantage? What if they plan to introduce a new technology with commercial advantage but new safety challenges? Transparency in safety analysis allows more thorough review and better stakeholder participation, but also gives competitors insight into near-term market strategy.

Dealing with Uncertainty

Related to reasonableness is the problem of limited knowledge. This may be scientific uncertainty related to new technology, or assurance deficits which could, if further resources were allocated, be reduced. The problem is that there will always be assurance deficits, i.e. knowledge gaps that prohibit complete understanding and perfect confidence (Hawkins et al. 2011]. No testing regime proves the absence of bugs, and no safety analysis demonstrates that all hazards have been identified and adequately treated. Organisational paralysis in the absence of perfect information would not be reasonable, but some level of doubt should lead to a pause and rethink of current operations. Indeed, Turner (Turner 1976) describes how many disasters arise not from willfully ignoring problems, but from a mistaken focus on the wrong areas of uncertainty.

Often it is not doubt itself, but the persistence of doubt which causes an ethical problem. When the first cases of blow-by occurred on the space shuttle solid rocket boosters, it may indeed have been appropriate to continue flying the shuttle whilst the problem was assessed and treated. By the time of the final Challenger launch, the *continued* uncertainty was a clear ethical failing. The lack of evidence that the O-rings would fail was a poor argument, since that evidence had not been properly sought. However, an engineer suggesting that "enough-is-enough" will almost always be challenged as to why they have not been more forceful earlier.

Their very willingness to be pragmatic and reasonable becomes ammunition against them when they finally take a stand.

Reconciling Competing Duties

We have already indicated that the four principles may intersect and conflict. Fulfilling one duty can mean compromising another. Even where the principles align, they may clash with other obligations which the engineer holds as a company employee and member of the wider community. For example, engineers not only owe a duty to support the interests of their employer, but any safety influence they have arises from meeting this duty. If a safety practitioner is not trusted to act in their employer's interest, and to act reasonably in balancing that interest with other concerns, they will not be able to meet their duty to the public safety.

A common escape from this apparent contradiction is to suggest that "safety is good business", but this is simply not true. Major accidents are fortunately rare, and in the short term even an organisation which neglects its own safety responsibility is still unlikely to have an accident. Organisations frequently gain short and medium term competitive advantage by ignoring rare, high consequence events (Taleb 2007).

Making Money

Safety management is an inter-organisational endeavour, shaped by contracts and informal business relationships. From a business perspective, the party responsible for performing safety work is the party being paid to perform the work. This viewpoint can be at odds with a "purely ethical" or legal determination of duties.

The Nimrod XV230 accident and subsequent Haddon-Cave Inquiry shows how ethical duties and contractual responsibilities can intersect in complex and counter-productive ways (Haddon-Cave 2009). The amount of work performed by the contractor was limited not by the amount of work *needed*, but by the size of the contract. The effectiveness of the independent assessor was limited by the ambiguous nature of their contractual responsibility, and by perceptions of their business motivations in raising safety concerns.

The business nature of safety presents obstacles to clear risk communication. Interaction between the parties is formalised, and practitioners often discharge their responsibility to accurately communicate risk through multiple layers of management and contracts. For example, it is clearly unethical to keep secret a potential hazard from those who will be exposed to it. An engineer who identifies the hazard must often rely on others to pass on their warnings. If management dilutes the message, or if the customer fails to respond appropriately, the engineer may not even know. If they *suspect* miscommunication, an attempt to investigate or to reinforce the message may breach duties of confidentiality and loyalty, particularly if the suspicion turns out to be unwarranted.

Ethics in Everyday Conduct

Ethical considerations can be present in some of the smallest everyday actions and decisions taken by engineers as they carry out their work. Firstly, there are many

issues around human communication and interaction in the conduct of safety engineering activity. For example, subtle choices can be made regarding the language used in statements in a safety case report that may dissuade, rather than encourage, challenge. Safety-related decisions can be heavily influenced by the personality types, and the interpersonal dynamics, of those involved on a project, as highlighted in (Haddon-Cave 2009). For example, the manner in which safety meetings are chaired can have a major influence on how safety issues are explored and sentenced. There are also practical considerations regarding the context and timing of safety engineering activities. For example, is it ethical for an engineer to review a safety report in the evening of an already busy day when they are tired? Should a safety manager schedule a safety review meeting when a known 'difficult' colleague is unavailable? Likewise, is it ethical to ask the lecturer on a safety course to provide endorsement of a complex safety decision within the space of a fifteen minute coffee break between lectures (a situation regularly faced by a number of the authors)? Equally, is it ethical for the lecturer to respond to such questions? Everyday choices such as these often have an ethical dimension and can have a significant effect on safety outcomes.

Choosing the Greater Good

The final challenge emerges from the nature of ethics itself. Different ethical systems can suggest different behaviour in the same situation. A set of ethical principles might suggest a deontological system - an engineer is ethical if they abide by the principles, trusting in the universal goodness of those principles[2]. However, what if the engineer could achieve a better outcome by breaking the principles? A consequentialist approach would suggest that the engineer should be concerned with the overall public good ahead of blind allegiance to principles[3]. The clash of these two approaches is not just a philosophical debate, but a grim reality for senior safety managers. "Choosing one's battles wisely" is the mantra of a consequentialist approach. How does a safety manager judge whether keeping their job is truly serving the greater good rather than self-interest? How does a safety engineer leaving their job reassure themselves that they really did all they could before forcing a final conflict?

4 Case Studies

In this section we discuss a number of case studies to bring the four principles and the challenges surrounding their day-to-day application into context. Producing a definitive set of answers to such scenarios is not feasible here. Instead, each scenario contains a short discussion highlighting the ethical concerns and relating these to the four principles.

[2] Although rule-consequentialists adhere to rules.

[3] But again it is not just consequentialists who give weight to consequences.

4.1 Case Study 1: Kudochem

The "Kudochem" study is adapted from the Royal Academy of Engineering's Engineering Ethics in Practice, specifically to illustrate the application of the Academy's third ethical principle 'Respect for Life, Law and Public Good' (RAEng 2011b). In this case the concept of respect for life involves "holding paramount the health and safety of others". The version of the case study below is an edited and reduced version to serve our present purposes.

Scenario
Kudochem is a multinational company producing chemicals for the agricultural industry. Responsibility for engineering issues at the eleven Kudochem plants in Europe lies with Kudochem's European Regional Engineering Director, Sally Proctor.

In the early hours of one morning, Sally receives a telephone call informing her that there has been a serious explosion at one of the plants. There have been some injuries, and damage has been done to property several hundred metres from the plant, but there have been no fatalities. The scale of the damage is huge, and the main site of the chemical plant is almost completely destroyed. In accordance with company policy an inquiry team is set up, involving company employees as well as independent consultants.

After several weeks, the team discovers two possible causes, both relating to a new ammonia production technique for fertiliser. This technique has recently been introduced in all of Kudochem's plants. The team is unable to determine which of the two possible causes are responsible. Given the presence of the production technique in all of Kudochem's plants, it is imperative that the ultimate cause of the explosion is identified, so that urgent steps can be taken to safeguard against similar accidents at other sites.

The inquiry team is very concerned at their inability to determine the precise cause of the accident. Without this knowledge, it will be impossible to satisfactorily modify the plants in order to prevent future explosions of this kind. They make a radical recommendation: to call a meeting with several competitor companies who are also using the new procedure in their fertilizer plants, in order to share experiences and research findings.

This would be a significant departure from standard practice, and some senior colleagues with commercial responsibilities have reservations. To call the meeting would entail releasing information about the safety lapse, as well as discussing sensitive commercial information with business rivals.

However, it may be the case that other engineers in other companies have encountered problems with the new method for producing ammonia, and could offer help in isolating the problem. Whilst such a course of action may be unusual in this case there are industries where safety critical information is routinely shared amongst competitors.

Discussion

In this scenario, the situation could be seen as one in which there is a conflict of interests and duties, such that Sally is required to balance these conflicting concerns. On the one hand she needs to ensure the safety of employees and local residents, and on the other hand she needs to maintain the security of commercially sensitive material. In addition, she needs to balance the risks with the financial costs of possible remedies, and she needs to judge what is appropriate in an abnormal situation.

The Statement of Ethical Principles states that an engineer must "hold paramount the health and safety of others." At the same time, though, she needs to take into consideration any other obligations she may have – including the duty to keep sensitive material secure, and to protect people's jobs by protecting the commercial interests of the company. Of course, if a company is acting illegally or irresponsibly, there may be a duty to 'blow the whistle', and this may defeat any obligation to keep sensitive information secret. However, in this case, there is no indication that the company was acting irresponsibly. As such, Sally could reasonably consider the commercial risks of sharing information with her competitors to be too significant.

Even if this was not her first response, she could be persuaded by commercial managers of the company that this is true. However, it is not clear that these considerations can outweigh the safety concerns. The principle states that she should *hold paramount* the health and safety of others. The same procedures are being used in all of Kudochem's plants and, given that the cause hasn't been identified, she needs to take seriously the possibility that there could be another explosion.

In summary, there does seem to be good reason to share safety information. Of course, where possible this should be done in a way that gives appropriate weight to one's other duties, regarding sensitive information, for example. Ultimately, however, it should be recognised that holding health and safety paramount doesn't just mean ensuring that you are not directly responsible for harms to the public, but that you also have some responsibility to help others improve their health and safety, for example by warning them of dangers they may not be aware of.

4.2 Case Study 2: Cargo Bay Doors

The Turkish Airlines Flight 981 crash was a real event. Knowing the future outcome of the decisions made in this case study gives the ethical considerations a prominence they would not necessarily have had at the time the decisions were made. This case study touches on all four principles, but particular emphasis can be placed on Principle 2 (honesty and integrity) and Principle 4 (responsible leadership). The main practical issues relate to the intersection of the principles with other duties.

Situation

Convair was responsible for the design and construction of the fuselage of the McDonnell Douglas DC-10. As a subcontractor to McDonnell Douglas, they undertook detailed design work, but requirements and major design choices were determined by McDonnell Douglas.

Most doors on an aircraft are of a "plug" design. They open inwards, and are held in place in flight by the pressure difference between the inside and outside of the aircraft. However, the cargo bay door on the DC-10 opened outwards; the pressure difference in flight pushed the door open, so it was important to have a reliable locking mechanism.

Convair conducted a hazard analysis of the door, which postulated several scenarios where failure of the door could lead to loss of the aircraft. They also had good reasons to question the reliability of the locking mechanism as failures had occurred in both ground flight and trials.

Convair were limited in their ability to directly control the safety risk of an open cargo bay door for several reasons:

- They supplied safety analysis to McDonnell Douglas and were contractually prohibited from speaking directly to the regulator. Safety issues that Convair raised to McDonnell Douglas were not always included in documentation that McDonnell Douglas passed on to the regulator, including the scenarios involving cargo bay doors opening in flight.
- The Federal Aviation Administration (FAA) were aware of the in-flight incident, and elected not to issue an Airworthiness Directive. They negotiated with McDonnell Douglas to issue a less enforceable Service Bulletin requiring minor changes to the door design.
- It was unclear who would bear the cost of changes to the door design, particularly if those changes were made at the request of Convair rather than their customer, McDonnell Douglas.

The Director of Product Engineering at Convair issued a memo to his immediate supervisor. In unambiguous terms he challenged the safety of the cargo bay door, and the adequacy of the changes made in response to the incidents. He predicted that at least one aircraft would be lost in-flight during the life of the DC-10. This prediction was fulfilled when Turkish Airlines Flight 981 crashed in France.

Discussion

The supervisor of the Director of Product Engineering at Convair had to weigh up the parties to whom he owed ethical responsibility. He was a subcontractor and under the regulation of the FAA he had to judge how these responsibilities were altered by the fact that his organisation had already made strenuous representations to McDonnell Douglas, and that the FAA had rules on the type and level of action in response to an earlier incident.

If the Convair director was not confident that his concerns had been clearly expressed to McDonnell Douglas, he would need to judge what further action, if any, he should take. He may have had uncertainty about his own interpretation of

what had already happened. Even if he was confident that his concerns had been passed on, he still had ethical issues to address depending on the feedback, or lack of feedback, he received.

Finally, this scenario also highlights the impact of contractual and financial arrangements on ethical decisions. Is the ethical concern about reasonableness affected by such arrangements?

4.3 Case Study 3: Independent Safety Advice

This case study concerns the ethical issues that a safety assessor may have to address when providing independent safety advice. It introduces issues relating to principles 2 (honesty and integrity) and 3 (respect for life, law and the public good). It also addresses issues of the importance of decisions and reasonableness.

Scenario

Swift, a Four Wheel Drive (4WD) passenger car manufacturer, has been approached by a large health organisation, Save, that is interested in expanding their rapid response vehicle fleet. Rapid response typically involves a multidisciplinary team attending to a patient often in very inhospitable environments. Quick delivery time of the fleet is crucial.

Save request that Swift involve an Independent Safety Assessor (ISA) in the safety approval process of the new fleet. Swift recruits Trust Limited, a safety assessment consultancy with a long established relationship with Swift. Trust has recently highlighted major safety concerns about the Electronic Stability Control systems of some Swift models. These concerns proved to be incorrect, resulting in expensive and unnecessary tests. Afterwards, Trust managers went to extraordinary lengths to keep Swift happy, leading to occasions where Trust ISAs were prepared to give undue credit during safety audit sessions to assurances made by Swift engineers.

Swift senior managers have a warm feeling that they can meet the safety requirements for the new fleet without making any significant changes to the existing vehicle models. A meeting was organised at the Swift Head Office and involved: Swift (senior business manager and chief engineer), Trust: (experienced ISA), Power (engine supplier powertrain specialist) and Save (director of regional operations).

At the start of the meeting, the senior business manager from Swift praised the safety track record of their vehicle product lines and supported his claims by the excellent feedback they received from their customers. The chief engineer highlighted that Swift complied with best practice and were annually audited by Trust. The powertrain specialist from Power agreed with Swift about their existing safety record but noted his concerns about the impact that the changes in the operational profile might have on the reliability of the engines.

The senior business manager from Swift dismissed these concerns, questioning the motivation behind them (i.e. “Power test people as usual are touting for business”). He asserted that it was always possible to carry out further tests but that this might exceed the budget allocated by Save, asking the powertrain specialist to be reasonable when making any such judgments. The chief engineer from Swift added that any claims about failures to meet the targets set by Save could only be made by the Swift engineers, who were ultimately the designers of the vehicle, and the engine was one of many vehicle components. The director of regional operations from Save explained that he was not an expert in vehicle safety and turned to the ISA for advice on the best course of action. How should the ISA respond?

Discussion

The Trust ISA is obliged to provide frank and honest advice within their domain of expertise. They are also ethically obliged to indicate if the advice that is being sought relates to knowledge or skills outside their domain of expertise. This is complicated by the fact that what appears to be an engineering judgement has an obvious direct impact on the negotiations between Swift, Power, and Save. The ISA is being employed by Swift, and has an implied duty to support them at the meeting. There are also a long term business relationships to consider.

This situation does not directly call into play principle 1 (accuracy and rigour) because the ISA is being asked to provide advice rather than information. Ethically, the ISA could act with honesty and integrity (principle 2) by declining to give an opinion, but would this show respect for life, law, and the public good (principle 3)? The Trust ISA has an opportunity to influence the safety of the vehicle, but their role is complicated by the contractual limits and financial perspectives that do not necessarily align with the boundaries of ethical responsibility.

If there are clear safety concerns, the ISA has a duty to highlight them. However, the ISA is not directly involved in the safety analysis work and as such their safety and domain knowledge is limited. Furthermore, because of the changed environmental and operational characteristics those involved cannot point to direct evidence, or counter-evidence, as to the safety characteristics the power unit and the vehicle will exhibit in operation. This is also complicated by the conflicting judgments made by the different stakeholders that are directly involved in the safety work. There is a trade off between the warm feeling of the designers and the cost of buying more information about the situation. This cost will be both financial and time. It is not clear who should bear the cost of such an information gathering exercise.

Uncertainty means that false alarms are bound to occur when addressing safety issues. The past consequences of such alarms will potentially have an impact on the confidence of the ISA to raise issues. It will also have an impact on the receiving organization trust in the advice given.

Finally, all parties should be conscious of behaviours that can compromise the integrity and independence of the ISA. For example, ambiguous communication and refutation of safety concerns, mixed with talk about business motivation, will

limit the effectiveness of the ISA. Prior agreement about the role of the ISA in this meeting could have made the ethical situation clearer for everyone.

Case Study 4: Safety and Novel Technologies

This case study concerns common ethical issues involved in the deployment of a novel technology. In particular, it focuses on principles 1 (accuracy and rigour) and 4 (responsible leadership), although there are aspects of the other principles.

Scenario

Robots R Us (RRU) specialises in developing safety mechanisms for automated cars. A notable success has been the development of a laser tracking system that is able to reliably model a 360 degree environment of the car with a visible radius of up to 150m. This will enable the vehicle to recognize actual and potential hazards as fast as a human driver, or possibly faster. In most cases the vehicle will either be able to avoid the hazard or to alert the driver in such a way that he or she will be able to take control in time to prevent an accident. The working name for the system is HATTAR (Hazard Avoidance Through Technologically Advanced Recognition).

RRU has a good relationship with a car manufacturer, Ovlo, with whom they have worked closely in developing HATTAR. While there is no contractual obligation to offer HATTAR to Ovlo, many at the company think that this would be 'the right thing to do'. Ovlo have also said that they would like to be kept up to date with developments on the HATTAR programme. However, there is also a concern that if the technology is sold to Ovlo, they will use it to gain a competitive advantage. Many feel that *all* automated cars should be as safe as possible and that no one company should benefit from a development that could ultimately save lives. Ovlo are working on their own safety system, which is not as effective as HATTAR. It is possible that they may take their own system to market first to recoup costs and then release HATTAR in five years' time in a new generation of vehicles. Every day in which this technology is not deployed people will die in accidents on the roads that could have been avoided by HATTAR.

A second concern is that the HATTAR technology has not been trialled on a large scale. There are foreseeable problems of HATTAR systems interfering with one another, leading to false positives in the recognition system. However, it is not realistically possible to carry out a trial on the scale necessary to fully test the technology. Ultimately, any hazards of HATTAR being used by a large number of cars simultaneously can only be determined through the actual use of the system on the roads. It will take a number of years until the HATTAR system is being used across the country in a manner such that its ultimate efficacy can be understood.

Discussion
This situation asks RRU to balance public good and private profit under high uncertainty. RRU could sell HATTAR to Ovlo and thereby pass the responsibility of its implementation and any inherent risks to the car manufacturer. This would likely mean restricting a powerful safety mechanism to one manufacturer. The HATTAR technology might remain in the hands of Ovlo and off the road for five years. On the positive side, the five years could be employed by Ovlo in further testing. Alternatively, RRU could refuse to sell the technology but continue to test it. With this option, RRU take the responsibility themselves for the resilience of the system. However, this will mean a long period, possibly several years, in which the system is not being used to improve road safety and a similar timeframe in which RRU will not benefit financially from this technology.

However, RRU could demonstrate industry leadership by making available the development to the entire car manufacturing industry. This would put the technology 'out there' to anyone who wanted to develop and test it further. This would have the greatest likely impact in terms of making automated vehicles safer. RRU would not stand to benefit financially from the revelation, though. To address this, RRA could patent and sell the technology to all manufacturers willing to pay for it. This has the advantage of making HATTAR widely available and benefitting RRU financially. However, RRU would lose control of testing the technology and the relationship with Olvo is likely to sour through RRU not having offered the HATTAR system to them first.

Whilst the principles offer little guidance to RRU in this situation, posing it as an ethical problem rather than a purely business decision allows RRU to properly evaluate their options with the public good as an explicit consideration.

5 Ways Forward and Practical Suggestions

In Section 2 we noted that the Royal Academy of Engineering has developed a set of ethical principles upon which engineers of safety critical systems can draw in order to develop good (ethical) judgement in their decision making. We then set out some of the challenges to implementing those principles in practice. Amongst other things, the principles intersect; they must be applied in complex circumstances with a range of other pressures bearing down; and they require a high level of interpretation. As shown by the case studies in Section 4, there is no algorithm for ethical decision making.

So how much use are the principles? Looking at matters more positively, we note that a very significant challenge for all of us is identifying an ethical problem *as* an ethical problem. The ability to perform that identification is a complex skill. Successful execution of the identification task depends on how we conceive or frame a situation. We can look at the same factual situation in very different ways.

Consider, by way of analogy, three people are on the same walk through the countryside. The first person might conceive of the situation as one in which they

are going bird-watching. If so, they will tend to notice the birds in their visual field. The second might see the same walk as field research into wild-flower conservation, and so notice the flowers. The third might conceive of it simply as physical exercise, and be oblivious to both birds and flowers. What the agent sees is affected by how they frame the situation. The Royal Academy's ethical principles provide guidance on framing engineering challenges so that the framing includes the relevant ethical considerations, not simply the technical issues.

How does one develop the skill of framing situations with the concepts set out in these principles? Simple awareness of the principles is a first step, but they also need to be made meaningful; as we noted above they are not self-interpreting. This is to some extent a matter of practice and experience – so working through the case studies above and others like them is a helpful technique. This quasi-experiential approach is related to the Aristotelian notion of habituation (Megone and Robinson 2001). Organisational culture can also help by finding ways to draw attention to the principles for staff.

Here are two related suggestions. One is to introduce an ethics template into the project planning process for all projects undertaken by an organisation[4]. The template is designed to help highlight the kinds of factors that engineers need to consider so that attention to ethical factors is built into the process of deliberation about project design.

Another process involves building up a database of case studies in which the Academy's key ethical principles are applied, so that in future deliberation the organisation can draw on past decisions. This allows them to consider how principles were weighed in different circumstances and to reflect on what that suggests for new situations. This process could start with a period of consideration on ways in which the principles are likely to apply to a particular organisation's safety critical engineering decisions. This may involve drawing up a provisional list of what seem to be the most significant concrete applications for that organisation. Then, on every subsequent major decision, it will be helpful to record how the principles were applied, as an audit trail of the approach to ethical decision-making. A third stage will then be to set a small amount of time aside, perhaps on an annual basis and possibly with an independent reviewer, to consider the application of the principles in the previous year, with a view to considering any particularly difficult decisions which might suggest that the organisation's guiding principles need refinement or revision.

This is a sensible practical process for embedding ethical principles in an organisation's practice, but it needs to be carefully organised (and perhaps involve some external review) in order to avoid the danger that it may just become a 'tick-box' exercise.

[4] Ethics Template, 2013, IDEA CETL, University of Leeds.

6 Conclusion

Every professional engineer is bound by a code of ethical standards. However, such codes do not provide a blueprint for all subsequent ethical action. In the engineering of safety-critical systems, there are a number of ethical challenges which arise in day-to-day work. These challenges include pragmatic considerations such as recognising when one is faced with the application of an ethical principle, and determining what is reasonable in communicating information about hazards. They extend to the philosophical, such as deciding what is ethical when the principles conflict, or even deciding whether a principles-based approach is applicable.

In this paper we have raised these issues and illustrated some of them through the use of four case studies. We argue that such reflection is itself a practical way forward for ethical practice of engineering. The principles can help engineers and engineering organisations recognise the ethical dimension of decisions, and by doing so, be more confident that they are resolving problems ethically.

We encourage engineers to reflect on their own practice, and organisations to provide safe spaces for discussion of ethical problems.

References

Dekker S (2009) Just Culture: Who Gets to Draw the Line?. Cognition, Technology & Work 11, no. 3

Haddon-Cave C (2009) The Nimrod review. The Stationary Office. London

Hawkins R, Kelly T, Knight J, Graydon P (2011) A New Approach to Creating Clear Safety Arguments. SSS '11, Southampton, 2011

HSE (2001) Reducing Risks, Protecting People, Health and Safety Executive Books

Lowrance W (1976) Of Acceptable Risk: Science and the Determination of Safety. Los Altos, William Kaufmann

Megone C, Robinson S (2001) Case Histories in Business Ethics. Routledge, London

RAEng (2011a) Statement of Ethical Principles. Royal Academy of Engineering

RAEng (2011b), Engineering Ethics in Practice: A Guide for Engineers. Royal Academy of Engineering

Taleb N (2007) The Black Swan: The Impact of the Highly Improbable. Random House

Thomson J (1976) Killing, Letting Die, and the Trolley Problem. 59 The Monist 204-17

Turner B (1976) The Organizational and Interorganizational Development of Disasters. Administrative Science Quarterly 21, no. 3