This is the published version of a paper published in Safety Science.

Citation for the original published paper (version of record):

Asplund, F. (2020)
Conflict as software levels diversify: Tactical elimination or strategic transformation of practice?
Safety Science, 126: 104682
https://doi.org/10.1016/j.ssci.2020.104682

Access to the published version may require subscription.

N.B. When citing this work, cite the original published paper.

Permanent link to this version:
http://urn.kb.se/resolve?urn=urn:nbn:se:kth:diva-269100
Conflict as software levels diversify: Tactical elimination or strategic transformation of practice?

Fredrik Asplund⁎, Greg Holland⁎, Saleh Odeh

⁎ KTH Royal Institute of Technology, Department of Machine Design, Division of Mechatronics, Brinellvägen 83, 10044 Stockholm, Sweden
⁎ Rolls-Royce plc, PO Box 31, Moor Lane, Derby DE24 8BJ, United Kingdom

A R T I C L E   I N F O
Keywords:
Communities of practice
Safety standards
Software levels
Standardization

A B S T R A C T
Communities of Practice create a shared consensus on practice. Standards defining software levels enable firms to diversify practice based on a software component’s contribution to potential failure conditions. When industrial trends increase the importance of lower software levels, there is a risk that the consensus on practice for software engineers used to primarily working at higher levels of assurance is eroded. This study investigates whether this might lead to conflict and – if so – where this conflict will materialize, what the nature of it is and what it implies for safety management.

A critical case study was conducted: 33 engineers were interviewed in two rounds. The study identified a disagreement between designers with different roles. Those involved in the day-to-day activities of software development advocated elimination of practice (dropping or doing parts less stringently), while those involved in expert advice and process planning suggested transforming practice (adopting realistic alternatives).

This study contributes to practice by showing that this conflict has different implications for firms that do not lead vs those that lead the early adoption of technology. At the majority of firms, safety management might need to support the organisation of informal opinion leaders to avoid vulnerability. At early adopters, crowdsourcing could provide much-needed help to refine the understanding of new practice. Across entire industries, crowdsourcing could also benefit entire engineering standardization processes. The study contributes to theory by showing how less prescriptive standardization in the context of engineering does not automatically shift rule-making towards allowing engineers to act more autonomously.

1. Introduction
Cyber-Physical Systems (CPS), enabling interaction with physical processes through information technology, have become well established in application domains such as healthcare, transportation, energy and manufacturing (Tomgren, et al., 2017). Despite efforts to simplify their engineering, entering the CPS market still requires expertise in a wide set of engineering disciplines (Geisberger, et al., 2015). Safety engineering is frequently one of these disciplines, since safety is often a critical characteristic of CPS. This expertise is codified in standards that directly or indirectly provide guidance on ensuring safety, such as DO-178C (aerospace software) (RTCA Inc., 2011), ISO 26262 (automotive) (International Organization for Standardization, ISO, 2011), IEC 60987 (nuclear hardware) (International Electrotechnical Commission, 2007), ECSS-Q-ST-40C (space safety assurance) (ESA-ESTEC, 2009) and EN 50129 (rail electronics) (CENELEC, 2003). As it is relatively costly to follow these standards, several of them allow for treating product components differently depending on their impact on safety. As an example, DO-178C recognizes and offers guidance for five software levels based on a software component’s contribution to potential failure conditions. The system safety assessment process determines the software level of a software component by identifying the level associated with the most severe failure condition to which it can contribute (RTCA., 2011). Software levels thus influence technological choices, both in regard to components and system architectures. However, the differences between levels are often found in the methods used in (or goals of) the development processes, rather than the technology employed.

In an engineering context, those implementing software according to software levels organise in various Communities of Practice (CoP), i.e. “groups of interdependent participants [that] provide the work context within which members construct both shared identities and the social context that helps those identities to be shared” (Brown and Duguid, 2001). CoP carry strong implications for their members: they...
introduce new practitioners to established practitioners and others related to their work (Lesser and Storck, 2001), and can develop their own pool of tacit knowledge and associated practices (Leonard and Sensiper, 1998). Software engineering practice thus involves alignment not only through mechanically adhering to processes, but through negotiation of meaning between those engaged in the practice (Wenger, 2010). In other words, the way work tasks are carried out will depend on what the collective of engineers engaged with them believe is required, to ensure the manufacture of safe products, for instance.

There are thus three efforts in the development process that establish the implications of the software level determination of a software component: firstly, the effort to handle the consequences of component failure identified during safety assessment; secondly, the redesign of the product to handle software levels efficiently in regard to both safety and cost; finally, the way work tasks are later influenced during implementation by how software levels are perceived by engineers. Of these, this paper focuses on the way work tasks are influenced by the perception of the engineers. More specifically, it focuses on the changes to this perception as technological trends increase the demand for software components at lower software levels. As the importance of levels that do not necessitate the strictest software assurance thus increase for development organisations, it can challenge the consensus on practice for software engineers used to primarily working at higher levels of assurance. This can lead to changes in the way CoP interact with each other and internally. Engineers might disagree on whether practice needs to change, and – if so – how it should be changed. Such disagreement can lead to a struggle, or conflict, for the right to decide on the future of practice within a firm. An important responsibility of safety management is to be pro-active in regard to such social interactions driven by change. Other organisational values than the wish to increase safety can drive practice, or at least be a source of resistance to such change (Grote, 2012). Unfortunately, standards that influence safety can currently be complex, open to interpretation and even ambiguous (Youn and Yi, 2014; Nair, 2014; Graydon, 2015). This means that even with a thorough understanding of a firm it could be difficult to deduce how firms should safely approach conflict regarding practice tied to different software levels. This motivates a case study into a context where the importance of lower software levels are increasing, challenging the consensus on practice for software engineers used to primarily working at higher levels of assurance. We explore conflicts due to this change and address three associated research questions:

Where do such conflicts materialize? What is the nature of these conflicts? Considering these conflicts, what are the implications of change for safety management?

The paper starts by constructing a framework from several academic discourses to serve as an analytical lens for the study. The next section describes our case as that of a large product to handle software levels e.g. the acting agency (Argyris et al., 1996). For the creation of shared meaning and handling power in relationships organisational learning through the adoption of relational practices to provide a social context allowing for knowledge sharing, can support edge-intensive organisations, CoP imply the existence of groups that, by developing by a whole organisational values such as safety.

Indeed, CoP are governed by the overall organisational values of the firms they belong to, but they might also develop their own specialized perspective on what is important in regard to work practices (Brown and Duguid, 2001; Leonard and Sensiper, 1998). Just as an entire inter-organisational CoP can extend across geographical and social boundaries in ways that are difficult to detect (Lave and Wenger, 1991), its intra-organisational parts can define a firm’s informal structure through which employees create, share and apply organisational knowledge (Lesser and Prusak, 2000). CoP can thus also be critical in facilitating organisational learning through the adoption of relational practices to allow the creation of shared meaning and handle power in relationships (Boreham and Morgan, 2004).

While an organisation’s processes describe the different steps that it takes to ensure safety during product development, its CoP thus provide the reasons why these processes were chosen in the first place. If communication issues between two different engineering disciplines have historically led to hazardous products, then both groups might have added separate precautions concerning e.g. product releases. This can be explicit in the processes, or implicit in whose words carry the most weight at meetings.

Much of the literature on CoP focuses on how they can be managed to the benefit of a firm (Bolisani and Scarso, 2014), a criticized shift away from the concept as an analytical one to an instrumental one (Wenger, 2010). In line with the original definition, the perspective here emphasises that members largely belong to CoP because they choose to identify with them. Members of CoP hold each other accountable to a common understanding of their community, establish norms of mutuality and have access to a shared repertoire of resources (Wenger, 2000) – thus defining knowledge and relationships that
establish what it means to be competent. Experiences by individual members outside of their CoP are brought back into the communities to change this collective understanding of competence. Organisational learning is thus dependent on the ability of individual members to suspend their identities as they cross boundaries between CoP, but also on their ability to negotiate the meaning of competency within their CoP (Wenger, 2000). Change can thus be difficult, especially if it challenges the currently dominant ideas on identity and practice (Roberts, 2006). Successful boundary spanning is a topic of research on its own, for instance in regard to gatekeepers who bring knowledge from external sources into an organisation and ensure that it is both understood and used (Paul and Whittam, 2010). The concept of CoP underlines findings in this discourse that suggest it is not possible to simply assign someone to the gatekeeping role, but rather requires identifying already well connected individuals and improving their networks and skill sets (Nochur and Allen, 1992). As the structure of CoP can be different from the formal structures at a firm, holding a formal position might not ensure that an individual is able to interact in a fruitful way with other experts. The different groups defined by formal and informal organisational structures can even be analysed from a political perspective (Tushman, 1977). The organisation of a firm into distinct parts can, for instance, create its own logic, leading to a self-reinforcing cycle where the flow of information is increasingly controlled (Vince, 2001). The CoP concept has been criticized for not taking such issues into account, focusing more on harmony and homogeneity (Wenger, 2010). This misses the point that informal structures centred on identity are inherently associated with power. One of the most basic implications for power in regard to CoP is that people tend towards homophilous communication, i.e. they favour communicating with those that are similar to themselves (Rogers, 1995). This stems from homophilous communication being more effective and perceived as more rewarding than heterophilous communication. Arguably it will be difficult to communicate new ideas between dissimilar CoP; since their differences imply both a high effort for cross-boundary communication and few connections between the groups. Trying to change practice as agreed by a community of practice might thus be a futile exercise if the right set of boundary spanners are not first persuaded to fall in line.

Not all suggestions on how to change practice that influences safety will thus have equal value. The details of best practice are defined by the practitioners working according to it. A well-known example is how management often form their own communities and might not consider the knowledge held by other CoP to be legitimate enough to merit consideration (Yanow, 2004). However, certain practitioners within a firm can, by force of their networking, be in a position to influence their community – especially if they are in a position to bring in and interpret knowledge from external sources. These experts can thus champion new ideas, and their existence explains how different CoP can influence each other – possibly leading to consensus between communities on cross-organisational issues such as safety. Such agreements between CoP could help identify the informal hierarchies within an organisation.

2.2. Software engineering

There are many high-level perspectives from which the practice of software development can be described, such as those based on knowledge areas (IEEE Computer Society, 2014) and sets of processes for building life-cycles (International Organization for Standardization, ISO, 2008). As the scope of these descriptions indicates, contemporary software engineering practice is large enough to encourage specialization into work functions focused on project management, design, testing, quality assessment, etc. Studies on knowledge management in software engineering have also highlighted the importance of tacit knowledge, that processes as they are performed can differ significantly from how they were formally designed, and the need for informal structures to support learning (Bjørnson and Dingsøyr, 2008). This implies the need for software engineering studies to take proper account of organisational roles in conjunction with their work practice. Unfortunately it is dubious whether this is usually the case. On the one hand, the academic discourse on software engineering has concerned itself mostly with technology and formal process, sometimes recognizing organisational roles but usually ignoring human and organisational factors (Lenberg et al., 2015). In some knowledge areas this is evident in the separate communities of academia and industry, with little evidence being generated by academia on several topics important to contemporary practice (Garousi and Müntylä, 2016; Garousi and Felderer, 2017). On the other hand, studies on individual characteristics frequently fail to consider the software engineering profession as heterogeneous, focusing rather on distinguishing between software engineering and other professions (Cruz et al., 2015; Beecham, 2008). The software engineering field also includes several high visibility concepts that may hide the importance of organisational and human factors, since they lead to different types of knowledge sharing (Ghobadi, 2015). As an example, agile development is acknowledged to rely more on the interpersonal skills of software engineers, blurring the line between organisational roles as loyalty to the team and organisation takes precedence (Ghobadi, 2015; Hoda et al., 2013). Arguably, more time has been spent on generalizing the difference between agile and traditional development than teasing out their different knowledge-sharing patterns and then generalizing the resulting implications.

The discourse thus suggests that CoP within software engineering can have important implications, but are understudied. This is likely due to software engineers being treated as a homogeneous group and the existence of other important phenomena that affect engineers and engineering processes.

The influence of different organisational roles has been noted, however – for instance in regard to how the difference in the practice of software development and testing may lead to communication difficulties (Zhang, 2014). Even though these groups work closely together they may thus form communities with different perspectives on each other’s skills and knowledge (Zhang, 2017). Interestingly enough, in the study by Hoda, Noble and Marshall on self-organizing teams in an agile development where organisational roles are transcended, testers are used both to make the point that the agile methodology needs to be reinforced to remain in place and as an example on how certain “personalities” might have to be removed in order to not hamper the agile methodology (Hoda et al., 2013). There is also evidence of software engineering managers forming a group distinct from other software engineers, for instance in regard to the possibility to codify knowledge (Dingsøyr and Røyrvik, 2003), the attribution of value (Taylor, 2016) and the importance of technical vs social skills (Kalliamvakou, 2017).

There are thus differences in the perspectives and practice of designers, testers and managers engaged in software engineering. This implies that CoP can emerge centred on these organisational roles. Similarly, what limited evidence there is concerning the importance of organisational culture in software engineering seems to carry implications that transcend individuals and development methodologies: Iivari and Huisman identify how a culture oriented towards stability and internal focus is related to the deployment of systems development methodologies (Iivari and Huisman, 2007); Iivari and Huisman, 2007); Iivari and and Iivari reflect on how organisations with a different orientation towards change vs. stability and internal vs. external focus have different implications for the efficiency of ad hoc agile and traditional development methods (Iivari and Iivari, 2011); and Siakas and Siakas suggest that agile development is suitable for organisations with a democratic culture (Siakas and Siakas, 2007).

Organisational values will thus have an effect on practice across organisational roles in software development. CoP within a firm cannot ignore its organisational values when defining software engineering practice. This suggests that the influence from the wider inter-organisational communities of designers, testers and managers will be interpreted locally in view of how much a firm emphasises safety.
2.3. Safety-relevant standards

Standards fill many roles in the development of advanced technological products, forming an infrastructure that has implications for both technology and economy (Tassey, 2015). Safety standards are of the type meant to specify acceptable product or service performance (Tassey, 2000), but are arguably unique due to their implications: failure to follow their guidance may not only impact the delivery of products but also the well-being of customers. While safety standards provide the same lowering of transaction costs as other standards that specify acceptable performance, they thus also have strong economic implications through liability: they provide a useful template for e.g. technical risk argumentation and best practice for avoiding hazardous errors (Kelly, 2014).

In other words, if a firm has adhered to the way a safety standard defines best practice, it may offer important protection should an accident occur. Safety standards are thus not only a strong moral argument for certain practice, but also an economic one.

However, the standards themselves are codified knowledge, which has to be absorbed and used by an organisation in order to have any impact. While not all safety standards define processes, it is arguably likely that an organisation will make use of safety standards by incorporating their knowledge in its processes. In this way the evidence required by a safety standard is generated as engineers perform their day-to-day duties. We note two implications of this approach related to knowledge handling. Most obviously it means that engineers might only be exposed to the bits and pieces of the safety evidence that relate to their processes. They may thus fail to appreciate the sum of what a safety standard strives for. Less obviously, as safety standards can be complex, open to interpretation or even ambiguous (Youn and Yi, 2014; Nair, 2014; Graydon, 2015), they often do not codify in complete detail the methods they stipulate. This can even be deliberate to allow alternative means of compliance. Internally to each organisation complying with a standard, there will thus be associated tacit knowledge known only to the engineers carrying out the associated methods. Software engineering organisations in particular rely on being able to efficiently share such tacit knowledge across the organisation (Bjørnson and Dingsøyr, 2008).

Engineers can thus form a selective view regarding the overall intent of the standards. Standards also only provide guidance to a certain level of detail. Ultimately, processes based on the same safety standards might thus be carried out in very different ways, as CoPs have leeway to interpret the guidance on safety.

Furthermore, safety standards frequently come in sets with other standards that are meant to be applied together (Youn and Yi, 2014; Nair, 2014). The “systems” aspect of standards, where several standards together create an overall impact on technology or economy (Tassey, 2000), is thus often relevant to safety standards. We use the term safety-relevant standards to indicate both safety standards and the standards they rely on to offer complete guidance. As an example, DO-178C for aerospace assumes that a set of complete, correct and consistent software requirements are given and offers detailed guidance on how to ensure that these are met (RTCA Inc., 2011). This means that DO-178C is typically applied together with ARP4754A (systems development) (SAE S-18, 2010) and ARP4761 (safety assessment) (SAE S-18, 1996). Combined, they form a large part of the recognized guidelines on how to structure the development of an aircraft to arrive at the evidence necessary for safety certification. Although this complexity and scope often means that safety-relevant standards are seen as incurring a high cost, some of the perceived cost could be avoidable through the use of contemporary software methods and tools (Youn and Yi, 2014; Wong et al., 2011).

Although standards such as DO-178C are not safety standards per se, software engineers thus often perceive them as such and they do indeed contain the guidance related to safety that is important to them. This “system” of standards might seem daunting and overly costly, but it could possibly be handled efficiently through contemporary or new practice.

Furthermore, engineers without deep expertise in applicable standards might over- or under-estimate the limitations these impose on development practices. This might not hamper engineers working with safety-relevant standards that prescribe the methods and processes to apply, as these often stipulate which techniques should be used at different software levels to establish confidence commensurate to the contribution of the software to system risk (Kelly, 2014). However, other, goal-based standards instead set out high level objectives for manufacturers to prove through the submission of a safety argument (Hawkins, 2013). Where goal-based standards are used, safety cases must explicitly justify the claim that the evidence it contains justifies such confidence, perhaps through the use of a separate confidence argument (Hawkins et al., 2011). Such justification can align with the guidance on levels, but will still have to argue for why the confidence established at a particular software level is appropriate. The likelihood of engineers influencing development practices negatively by over- or under-estimating the limitations that safety-relevant standards impose on them is then aggravated by two aspects of these standards: firstly, that the process for creating standards excludes relevant experts from learning through participation (Habli, 2017); secondly, that the rationale behind many safety-relevant standards is implicit (Habli, 2017). Engineering firms will most likely come under increased pressure to handle these aspects to achieve an increased understanding of the rationale underpinning the guidance provided by safety-relevant standards.

In other words, if standardization moves towards goal-based standards in the future, it increases the chance that any argument for the use of software levels will be required to be explicit. This also increases the pressure on engineering firms to establish ways through which their engineers can learn the underlying intent behind the safety-relevant standards they use.

2.4. Analytical framework

This subsection summarizes the most important findings presented so far. It provides a succinct base for understanding our approach to generating and analysing results.

Our analytical framework focuses on the primary CoP that we can expect to encounter in association with software development, i.e. software designers, testers and managers. Internally to firms, members of each community form a consensus on best practice through organisational learning primarily under the influence of two sources: that of the internal informal power structure and that of the entire external community that spans the software development industries. Standards important to organisational values, such as safety-relevant ones, have implications for both the internal and external paths of influence on consensus.

Internally to an organisation, the interactions between members of the primary CoP will lead to a partial overlap of perspectives reflecting the informal power structure and boundary spanning in each firm. An intra-organisational community that holds significant informal power will export its perspectives to other intra-organisational communities. As safety-relevant standards are perceived as codified best practice, members of an intra-organisational community can use them as moral and economic motivation when struggling to change the consensus on best practice according to their perspective. In other words, in the process of organisational learning of practice within a firm, standards are leverage in conflicts between communities.

Externally to an organisation, CoP continue evolving their definition of best practice. As new practices emerge and become increasingly accepted, this can lead to new means for complying with a standard becoming seen as a realistic alternative by practitioners. Members of an intra-organisational community can thus attempt to change the consensus on practice by importing realistic alternatives to it. During this
process, those struggling to resist change can point to how a new practice is not codified in a standard, while those struggling to change can point at the acceptance of the practice and that the standards allow for it. In other words, in the negotiation of practice within communities, standards can also be leverage.

Fig. 2 summarises our analytical framework.

3. Research design and methods

This section provides a case description, positions the research as a critical case study, and details the data collection, data analysis and validation of the study.

3.1. Case description

The firm on which the study is based is a multi-national engineering company developing CPS. The firm employs about 50,000 employees in 150 countries developing products for both civil and defence purposes in the aerospace, marine, nuclear and power domains. Engineering is set up in organisationally separate business sectors focused on the different domains. However, certain capability functions and initiatives have an enterprise-wide reach. The CoP concept is actively supported by the firm, with more than 350 communities currently organised in a bottom-up way around shared interests. These interests include design and verification, partitioned across disciplines and domains. Furthermore, in line with the original view of CoP as groups self-organising a mutual learning process, the firm also operates several extensive initiatives for bottom-up knowledge sharing – including an innovation portal, a social media platform and several wikis related to engineering.

This study refers primarily to the civil aerospace part of the business. Being the largest sector in the firm with about 20,000 staff directly and indirectly involved, it has existed in various forms since the inception of the aerospace domain. The software experts within the aerospace sector have an in-depth understanding of applicable safety-relevant standards and have contributed actively to standards such as DO-178C, ARP4754A and ARP4761 throughout the last three decades. They are also active in working groups relevant to many aspects of software engineering, such as those of the Object Management Group, the International Council on Systems Engineering and the United Kingdom Safety-Critical Systems Club. Furthermore, the primary software unit in this business sector includes a group dedicated to supporting software engineering across all business sectors, for instance through knowledge transfer between them. Strategic outlook capabilities thus include both a depth in the aerospace domain and a breadth across several other domains.

Since the software level concept was incorporated into DO-178A (RTCA, 1985), the business sector has returned to it at each major update of product architectures. As the focus of the software development organisation in this business sector is on software components at the highest level of assurance, the approach to diversification into software levels has traditionally been to limit it. However, software is still currently being developed in accordance with several of the levels defined by DO-178C.

This development of highly safety-critical systems is currently changing due to influences from the wider CPS industry, where the use of smart functionality such as Artificial Intelligence (AI) and predictive analytics is increasing quickly (Törngren et al., 2017; Geisberger and Broy, 2015). The technical impact of these trends is straight-forward. Manufacturers make use of smart functionality. However, they do not make use of all functionality made feasible by such technology, as some of it would imply that the associated components would have to be assured at a high software level. The software level of the associated components is thus kept low, allowing manufacturers to avoid the
The interviews were intended to both verify the existence of a phenomenon and to change it (Runeson, 2012). That case studies can be used to explore or describe a phenomenon is generally accepted, while the explanatory power of case studies is a more contentious issue (Runeson and Höst, 2009). This is due to explanatory research in case studies typically relying on a qualitative understanding of a phenomenon in its context to argue for the generalizability of conclusions (Runeson, 2012).

The studied firm supports standards both existing and under development, interacts with organisations gathering inter-organisational CoP, and provides intra-organisational support for the CoP concept. The firms’ CoP thus enjoy the freedom to organise bottom-up and hold significant informal power, while receiving direct access to current, evolving and future best practice. With the business sector in focus comes a large international stable environment which has revisited the concept of software levels over several decades. This case thus forms a critical case (Flyvbjerg, 2011): the CoP can be expected to strongly set the consensus and agenda on practice with little preconception of what constitutes valid practice; if disagreement regarding how to change practice has never existed or led to significant struggle, then one should not expect anything but case-specific conflicts in other firms. Using our case we can thus with strong confidence verify the existence of any underlying generalizable conflict centred on software levels within or between CoP. This case is also ideal for exploring the characteristics of such a conflict, since engineers at the firm are likely to have experienced it over time. However, by being a “most likely” critical case (Flyvbjerg, 2011), it is not well suited to exploring other conflicts that might only manifest under specific circumstances. Therefore, this is not an objective of this study.

3.3. Data collection

Data was collected through semi-structured interviews according to the procedure defined by Brinkmann and Kvale (Brinkmann and Kvale, 2015), which includes thematising and designing the investigation; conducting, transcribing and analysing the interviews; and verifying and reporting the results.

The thematising of the interviews started when the case and research questions were elicited as part of the planning of the case study. The interviews were intended to both verify the existence of conflict and describe it, with the case allowing for generalizable results. To ensure that findings were not simply a reflection of prevalent organisational values, it was decided that these should first be identified through a preparatory round of interviews conducted across the firm. The sampling for this preparatory round was opportunistic, focusing on known opinion leaders across the firm, and ultimately included 13 people. The selection of interviewees for the primary round was by contrast careful, focusing on covering a representative and knowledgeable sample from each of the three primary CoP. The characteristics of the resulting sample of 20 people are reported in Table 1. As part of the data analysis, we identified two groups of designers: signers with a tactical role, i.e. those involved in the day-to-day activities of software development; and designers with a strategic role, i.e. those involved in such things as expert advice and process planning. The last column in Table 1 lists to which group each software designer among the interviewees belonged.

An interview script was designed for each round and community. These scripts ensured that all topics were covered. They also allowed for the interviewers to “push forward” (Brinkmann and Kvale, 2015), as several follow-up questions were explicitly noted in the scripts. This ensured that researchers were reminded to clarify the meaning of ambiguous statements. The threat to internal validity of interviewees providing biased or incorrect information was also considered. This was deemed most likely if a response could somehow affect an interviewee’s career negatively. Therefore, all interviewees were assured that no data would be shared from the interviews that could be used to identify an interviewee.

Interviewers need to listen actively (Brinkmann and Kvale, 2015), otherwise internal and construct validity can be compromised by ambiguous responses or failure to follow up on important leads. To ensure active listening in this study, at least two interviewers, sometimes three, were present at each interview, each taking turns to either ensure that the interview script was followed or focus on the interviewee’s responses. For the primary round a pre-interview questionnaire was also sent out, which supported active listening through the opportunity for the interviewers to discuss the script in more detail prior to the start of each interview. Each interview took about 1 h to conduct and all were conducted face-to-face.

All interviews were recorded and then transcribed by the firm’s transcription service. To ensure the reliability of the data the transcribers were instructed to leave parts that were difficult to transcribe to the interviewers. As the aim was solely to capture the meaning of the interviewee’s comments, they were not transcribed verbatim. However, neither were grammatical errors corrected. This ensured that the data analysis could identify ambiguities and refer back to the recordings in order to handle them.

3.4. Data analysis

The data collection for the preparatory round took 1 month, followed by 3 months of analysis and verification. The data collection for the primary round took 2 months, followed by 6 months of analysis and verification.

During the initial analysis of each round, each interviewer coded the transcriptions with descriptive codes (Saldaña, 2009). All three interviewers then met weekly to discuss the codes and arrive at a common code book. The code book for the preparatory round eventually included 167 codes, and the code book for the primary round 174 descriptive codes. All interviewers agreed that the meaning and application of these codes were consistent. This was required to ensure that internal validity was not affected by unreliability of the coder or coding.

The initial analysis was followed by weekly meetings focusing on recoding the descriptive codes to identify patterns (Saldaña, 2009). The resulting secondary coding aimed at interpreting the meaning of the interviews in light of the analytical framework (Brinkmann and Kvale, 2015). Through this process, patterns were developed iteratively, which resulted in the themes reported in Section 4. This ensured that the interpretation was free of contradictions, even when the findings seemed paradoxical at first glance. It also ensured that the overall interpretation could be tested against its parts, the firm and the literature forming the analytical framework – all important parts of analysing the meaning of interviews (Brinkmann and Kvale, 2015). An example of the development of the categories is given in Fig. 3 to help clarify the process to the reader and to illustrate how initially contradictory statements were reconciled.

3.5. Validation

As outlined in the previous subsections several actions were taken to ensure the internal validity, construct validity and reliability of the study: the preparatory round ensured that results were not simply a reflection of the firm’s organisational values; the design of the interview script, the pre-interview questionnaire and the use of several
Table 1
Profiles of the interviewees from the primary round.

<table>
<thead>
<tr>
<th>Interviewee</th>
<th>Experience at Firm (years)</th>
<th>Background at Firm</th>
<th>Community of Practice</th>
<th>Tactical or strategical designer</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>20–25</td>
<td>Design</td>
<td>Design</td>
<td>Strategic</td>
</tr>
<tr>
<td>2</td>
<td>20–25</td>
<td>Design</td>
<td>Manager</td>
<td>(Not applicable)</td>
</tr>
<tr>
<td>3</td>
<td>20–25</td>
<td>Quality Management</td>
<td>Manager</td>
<td>(Not applicable)</td>
</tr>
<tr>
<td>4</td>
<td>20–25</td>
<td>Design and Verification</td>
<td>Verification</td>
<td>(Not applicable)</td>
</tr>
<tr>
<td>5</td>
<td>20–25</td>
<td>Verification</td>
<td>Verification</td>
<td>(Not applicable)</td>
</tr>
<tr>
<td>6</td>
<td>20–25</td>
<td>Design and Verification</td>
<td>Verification</td>
<td>(Not applicable)</td>
</tr>
<tr>
<td>7</td>
<td>20–25</td>
<td>Design and Verification</td>
<td>Manager</td>
<td>(Not applicable)</td>
</tr>
<tr>
<td>8</td>
<td>20–25</td>
<td>Design and Verification</td>
<td>Design</td>
<td>Strategic</td>
</tr>
<tr>
<td>9</td>
<td>20–25</td>
<td>Design</td>
<td>Design</td>
<td>Tactical</td>
</tr>
<tr>
<td>10</td>
<td>20–25</td>
<td>Design</td>
<td>Design</td>
<td>Strategic</td>
</tr>
<tr>
<td>11</td>
<td>20–25</td>
<td>Design</td>
<td>Design</td>
<td>Strategic</td>
</tr>
<tr>
<td>12</td>
<td>20–25</td>
<td>Design and Verification</td>
<td>Design</td>
<td>Tactical</td>
</tr>
<tr>
<td>13</td>
<td>15–20</td>
<td>Design</td>
<td>Design</td>
<td>Strategic</td>
</tr>
<tr>
<td>14</td>
<td>20–25</td>
<td>Programme Management</td>
<td>Manager</td>
<td>(Not applicable)</td>
</tr>
<tr>
<td>15</td>
<td>20–25</td>
<td>Design</td>
<td>Design</td>
<td>Tactical</td>
</tr>
<tr>
<td>16</td>
<td>15–20</td>
<td>Verification</td>
<td>Verification</td>
<td>(Not applicable)</td>
</tr>
<tr>
<td>17</td>
<td>20–25</td>
<td>Design</td>
<td>Design</td>
<td>Strategic</td>
</tr>
<tr>
<td>18</td>
<td>20–25</td>
<td>Design</td>
<td>Design</td>
<td>Strategic</td>
</tr>
<tr>
<td>19</td>
<td>20–25</td>
<td>Verification</td>
<td>Verification</td>
<td>(Not applicable)</td>
</tr>
<tr>
<td>20</td>
<td>15–20</td>
<td>Design</td>
<td>Design</td>
<td>Strategic</td>
</tr>
</tbody>
</table>

Fig. 3. Pattern codes to descriptive codes.

Interviewers removed ambiguity; the sample of interviewees ensured a complete coverage of perspectives across the firm; anonymity minimized the risk of false or biased information; several interviewers decreased the chance of interviewer bias; following up on uncertain transcriptions increased the reliability of data for analysis; analysing and coding together meant coder bias was minimized and coding reliable; and testing interpretations against each other, the firm and the organisational framework ensured consistency.

The cooperation on preparing the interview script, coding and analysis was a major part of ensuring internal and construct validity by minimizing bias, as the researchers all came from different backgrounds. Member checks were also used to ensure the internal and construct validity of the results (Creswell and Miller, 2000). This meant that interpretations and conclusions were continuously checked with employees of the firm, and the complete study eventually presented and verified by two of the interviewees as well as other senior employees at the firm.

The external validity of the study is primarily based on the analytical generalisation presented in Section 3.2 (Brinkmann and Kvale, 2015). Identifying underlying generalizable conflict should reasonably be within the abilities of this study. However, it is difficult for this study to define the magnitude of the effect of such conflict at other firms, since these may exhibit other case-specific conflicts that counter or enhance the effect. This is not unexpected in qualitative studies, as the transferability of results is often a prerogative of the reader (Brinkmann and Kvale, 2015). Further research is thus required to verify any implications of the results in smaller, less international, more hierarchical or less externally facing firms.

4. Results

This section presents the results from the study in the form of themes arising from the interviews. These are grouped into tables based on which CoP shows consensus on the theme, with example quotations to clarify their individual meaning. Summaries are provided after each table to clarify the combined meaning of each group. When deemed appropriate to simplify reading, examples from the tables have also been reproduced together with the summaries. Similarly, when quotes have been obfuscated to hide the interviewees’ identities, some examples are also provided together with the summaries.

4.1. All in agreement

Tables 2, 3 and 4 group those themes on which all in the firm agreed, including all members of the primary CoP pursued. Interviewees across the board stated that they had a high awareness of the importance of individual employees acting in a professional manner, especially with regard to the possible implications of the firm’s products on safety.

Everyone also stated that the firm’s comprehensive standard-based process descriptions were not the primary factor behind engineers adapting a professional safety-minded engineering approach. According to the interviewees this practice was primarily learnt by engineers working together with other engineers.

The interviewees stated that cost was the largest issue with working at higher assurance levels, and the reason firms would want to develop software components to lower levels. However, according to the
Table 2
Example quotations on core values.

| Professional responsibility | “… they should have a responsibility for doing the best they can and taking, yeah, I don’t know, whatever you want to call it, ethics and environmental and all those other things that say well actually I’ve got the – and obviously responsibility to the company as a whole.” |
| Safety as a Core Value | “So your professional responsibility to [Firm] is to deliver a quality product that’s safe that meets the customer requirements, is efficient, etc., the whole raft of different criteria around it. You’ve got to take professional responsibility for delivering that.” |

Because I think most engineers realise – well certainly if you’re working in this industry that what you do needs to be right because of, you know, if you make a mistake that can, you know, could be catastrophic. … and there’s just the general sense of kind of pride in what you do, that you want to be professional, you know, you are a professional person and therefore you want to behave in a professional manner, through your own sense of self-worth really.

“…the way I’m working, there isn’t a distinction in terms of what I’m doing, you know, what I hope is that, I’m conscious of what the appropriate level of safety is, of what the right level of rigour is that’s needed to support that, and that I’m applying that. So in terms of being conscious of the safety consequences, I would say I am, and I believe working to the right levels of doing the right activities for that.”

“So almost my, you know, my involvement in that, so I ultimately feel very accountable for safety…”

“Process doesn’t make me feel that way [accountable]. Me doing my job to the best of my ability makes me feel that way.”

“Engineering judgement is – and a sort of moral value to say, I’ll speak up when I need to. I mean I’d think that’s professionalism to be honest.”

“Because if you don’t take responsibility you’ve always got at the back of your mind not only the monetary side but more importantly the safety of all the people out there. And a slight mistake on the design or the manufacture of a part in our [product] could lead to a disaster.”

“... every year or two we have to do this online learning about all the, who’s responsible- I did it, just did it recently again, this online learning stuff, so there is, there is commitment to sort of product safety …”

“… you know, we’ve got a duty to put out a product which is safe and just – in my mind just meeting the regulations is a small part of that. You know, you have to think more widely about, you know, what – practically how can we reduce the risk as low as we can?”

“Well, we had an issue on [CPS Product] some years ago and we had a very focused kind of team all the way from the top to the bottom addressing that particular incident.”

“I think cost and trying to reduce cost comes kind of second to that … we’re always, absolutely a hundred-percent committed to safety, I don’t think that’s a primary thing that is leading people to make decisions differently.”

“Safety always take priority at the end, there’s no problem with that.”

“I think it’s the enormity of the product and the job we do. We know people are [using CPS Product], we know that they count on everything that we do … We have a lot of training on safety, and a lot of that is focused on just getting you to wise up about the seriousness of what we do.”

Table 3
Example quotations on learning.

| Learning from Others | “I don’t think it’s credible to give a newcomer a thirty page process, expect them to then follow it. Because by the time they get to page five they’ll have forgotten page two. So at which point they’re going to get the gist of the process and ask the guy next to them, what do I do? ”

“… you’ll read the process and then see how people were using it, but you would follow how people were using it.”

“You’ll ask somebody that you’re working with … I think it’s human nature to say, how do I do this and somebody will show you what they do.”

interviewees the current practice, associated with higher levels was very much ingrained in the engineering workforce and unlikely to change to accommodate lower costs when moving to lower levels.

4.2. Agreement across CoP

Table 5 groups those themes on which software designers and testers agreed, and Table 6 gives the themes on which software designers and managers agreed. This provides additional detail to the themes identified across the firm.

Designers and testers attested clearly to the importance of standards and processes, but maintained that there was little need for the average engineer to frequently refer back to them to ensure their correct use. Indeed, they stated that there were other, larger risks associated with well-structured processes divided into multiple independent steps: the processes could pigeonhole employees, indirectly decreasing the employees’ ability to maintain product safety by decreasing the understanding of how activities ultimately contribute towards this goal. Furthermore, the independence could lure engineers into a sense of security based on later process steps, directly undermining product safety by decreasing the rigour of engineering activities. Designers and testers stated that avoiding these issues was beyond the abilities of the average manager, who relies on engineers to provide the appropriate understanding of practice to get the processes right.

Designers and managers both stated that there was a need to change some parts of the current practice to work cost-efficiently at multiple software levels. Simultaneously, rather than eliminating as much practice as possible at lower levels, they agreed that much of the other parts of best practice should be kept even if this required a large costly effort. Among these other parts of best practice, several of the interviewees mentioned the need to keep performing reviews of all aspects of high-level requirements, and the importance of keeping these reviews independent of those creating the requirements.

4.3. The design community – tactical vs strategic

Engineers appeared to wield informal power through their specialist knowledge, as the perspective of software designers was a common denominator across the primary CoP. This suggested that we take a closer look at this community, which – as previously mentioned – led us to identify two groups: designers with a tactical role, i.e. those involved in the day-to-day activities of software development; and designers with a strategic role, i.e. those involved in such things as expert advice and process planning. We discovered that these groups held partly conflicting views, but more importantly that they each had detailed explanations for the themes described in previous subsections. Table 7 reports on the resulting categories.

On the one hand, tactical designers stated that dropping some parts of existing practice when moving to lower software levels was possible, as long as a proper assessment was performed. At the very least there was ample opportunity to perform some practice less stringently under these circumstances. Examples of such changes to practice involved working in pairs to improve the design process and ensure quicker feedback, while at the same time dropping the independence between the developer and the reviewer of a work artefact. Another example was to increase the number and scope of changes addressed by change management during a specified time period. This would increase the risk that a fault would be overlooked. However, assessment would have indicated that such faults would be identified later by other means, and this change would also allow quality improvements to find their way...
Table 4
Example quotations on changing the assurance target.

<table>
<thead>
<tr>
<th>Cost is the Reason</th>
<th>Old Ways Ingrained</th>
</tr>
</thead>
<tbody>
<tr>
<td>“Well, I think for a number of reasons. Cost is very important. [Highest assurance] is very, you know, verification-intensive and verification being a large part of the cost.”</td>
<td>“I mean, there’s resistance to change normally anyway, and for people who take pride in doing a complete job because they’re aware of the consequences of something going wrong, without explaining to them properly that this is the reason why we’re now behaving like this and we’re able to take such-and-such a step out because it’s either not valuable or we’ve changed the structure of the system so that it doesn’t matter so much if this component fails or this software feature fails, or we’re able to take the risk, no one’s going to die because we haven’t done this level of testing if actually something untoward is there and we haven’t found it, I think there’d be resistance.”</td>
</tr>
<tr>
<td>“I’ll say the main reason I can see why a business would do it [work to lower assurance], it would be to try and save money.”</td>
<td>“When we do [lower assurance] projects we typically just routinely just do [Practice A] anyway, because that’s how all – that’s our culture, that’s how engineers are brought up …”</td>
</tr>
<tr>
<td>“Obviously because of the cost, [highest assurance] is expensive.”</td>
<td>“People think [highest assurance requirements] is the only means of getting correct software.”</td>
</tr>
<tr>
<td>“The major reason, of course, has to be with cost because [highest assurance] of software are certainly perceived as being vastly more expensive …”</td>
<td>“Yeah, I’d say that most people are, you know, it’s a very – from the point of view of people taking safety seriously, [software department] do that exceedingly well. Everyone is concerned with safety. They look at it all the time. But the point is they tend to like to work in one way, therefore they’re always working at the [high assurance] end of things all the time. That’s their mentality. And I can understand that.”</td>
</tr>
<tr>
<td>“There are requirements in [Standard] that mean that [highest assurance] are almost certainly going to be more expensive. So one answer would be, in their pursuit of cost reduction …”</td>
<td>“You can never go back – you can’t take a nervous engineer who’s been trained in a culture of [high assurance] software to suddenly cut back on their standards. That’s going to be very tricky.”</td>
</tr>
</tbody>
</table>

quicker into the product. Tactical designers stated that this would rely on engineering judgement and informal coaching of newcomers, as they could only identify an implicit link between development activities and product risk. On the other hand, strategic designers stated that they were negative to such attempts, and for several reasons: it could incur large costs later in the life-cycle of components; there were several examples of how dropping some parts of practice had had a negative effect on development; and large reduction in cost would not come by dropping some parts of current practice, but rather with the introduction of new practice to support product functionality currently unfeasible to introduce at the highest level of assurance. The strategic designers gave several examples of such new practice. One example was the use of Commercial-off-the-Shelf components and techniques for ensuring that these complied with written specifications or given performance requirements. Another example was the use of algorithms for machine learning, specifically when one could take advantage of certain use cases not requiring these algorithms to be deterministic. None of these examples were mature solutions, but rather early suggestions based on an appreciation for the challenges of contemporary safety-critical systems development. They all either sought to decrease the reliance on or change the character of the required process evidence at the highest level of assurance. Instead, these suggestions attempted to leverage on (novel) characteristics of the software components and the product environment to ensure confidence commensurate to the contribution of the software to system risk.

5. Analysis

This section addresses the first two research questions by analysing the results. It is thus our interpretation of the interviewees’ responses in Section 4 based on the analytical framework provided in the end of Section 2. However, where appropriate we refer back directly to the framework or the interviewees’ statements as summarised by Table 2–7.

5.1. A conflict on the implications of diversifying

The importance of organisational values and informal learning at the firm, as described in Tables 2 and 3, arguably lines up well with the CoP concept. Superficially, the situation seems straightforward to explain based on the themes elicited across all interviewees. As exemplified in Table 4, the general perspective by the interviewees was that working primarily at the most rigorous software levels will ingrain the practice required at these levels into the organisational culture of large firms. While firms can, to lower costs, try to drop part of this practice by diversifying development into several software levels, they would struggle to reap the associated benefits. Engineers used to working at the highest software levels would not be comfortable with changing their practices even when the implications of a component failure is not severe. If the interviewees are correct the result would rather be a conflict on practice between management and other engineers, with management at a disadvantage due to the specialist knowledge of designers and testers. The effect of this informal power is also seen at the firm: in Table 5 designers and testers describe the importance of their expertise, and in Table 6 managers echo the discussion within the designer community. From this perspective the evolution of product architectures provides opportunities to accommodate smart functionality, but engineering practice inertia acts as an obstacle to realising these opportunities.

We argue that this type of conflict can be described as centred on single-loop learning in line with the internal path to influence consensus outlined by the analytical framework. The different CoP observe an unintended consequence in the too high cost of development, but managers and designers disagree on how to address it. However, due to the implicit nature and technical complexity of the associated safety-relevant standards this disagreement is arguably unlikely to lead to a direct struggle for the right to decide on the future of practice within the firm. As designers are in possession of the skill required to interpret the standards, they have the final say on whether practice will change or not. The situation could of course be altered if safety became less important as an organisational value at the firm, or if managers recruited engineers from outside the safety-critical industry to develop components to lower levels of assurance. Strictly speaking, the former case has implications far beyond that of software levels and is best studied separately – for instance in relation to firms close to economic collapse. The latter case could change the situation for a time, but the designers would still be interpreting the standards. The newly employed designers could be kept separate from the rest of their community, but eventually the common connections to safety management and the importance of safety should lead to interactions. The discussion regarding the future of practice at the firm would then occur inside the designer community.

The example statements in Table 7 from the designer community by contrast suggest how disagreements could lead to conflict, as perspectives differ between tactical and strategic designers.

To explain the situation one should note that both of these groups will agree on the lowest possible software level for any given software component and product architecture, as it is decided by the product’s current safety assessment. The divergence arguably lies in their roles leading to different ideas about why and how to change practice to ensure that this development is sustainable. Tactical designers are
Table 5
Example quotations on agreement between designers and testers.

| Process Pigeonholes People | “I think so because I think we're kind of — certainly now we're geared up to very big teams and people just doing a very tiny part of a job. It just seems to have grown up over the years that teams have become bigger and bigger and more complex and people are isolated.” |
| Process Makes Me Feel Safe | “They work in that bubble and that's what they do. So it'd probably depend on where you are in your career and what's gone on. So early in your career, and some people just stay there, you will be pigeon holed into you're doing a certain activity and you're doing it for the next six months. Some people stay there forever.” |
| Power to the Specialist | “I mean if you're just doing something, which most people do, in isolation... well, it's not necessarily going to be evident that if you get your bit wrong... I'm sure it will be because some people won't have an overview of the whole system because they've never actually done anything with the whole system; they've just done... I wouldn't say always the same part, but just parts of that they can get to put together to make the system.” |

“Process Pigeonholes People”

“Process Makes Me Feel Safe”

“Power to the Specialist”

“Infrequent and Partial Reference to Standards and Processes”

“I should imagine that for most test engineers in [Department] they have a specific [Department] test process, they read it when they first start, it's not very complicated, it tells us what to do. Every now and again we look at it, if it's a bit of clarity, probably once every two or three years, just to re-familiarise myself.”

“What I have done in more recent years is look at specific elements of that, so things like review checklists for example, I've not looked through all of the process every time, what I've thought is, I understand what we're doing, I understand the basic process, every now and again I think, I'm not quite sure I've remembered exactly all the details of a certain part of it, so I've gone to look for those. That tends to be what makes me go to look.”

“Very few, because most people have to just do one point in the process many times. So they might have to produce a low level test, do six months at low level test, so they'll go and ask the person who knows how to do level tests on that project how to do it and that's probably adequate. The team leads and people like myself might step back to the plan set, and work out way through from that. If I'm doing a stage of involvement audit I'm actually not looking for process compliance anyway, I'm looking for certification compliance. So I'm looking that the objectives of [Standard X] have been met or [Standard Y] or [Standard Z] have been met, so I don't need to look at the process reviews to produce the stuff. I only need to look at the stuff. So yeah, we plan that, we write it. Once it's been socialised, once it's embedded that's how we do stuff we don't go back and refer to it every day to work out what to do.”

 pressured to minimize the risk of not delivering on time with the available resources. This risk is minimized by keeping to well-known practice, but lowering the constraints on software development when available resources. This risk is minimized by keeping to well-known practice, but lowering the constraints on software development when available resources. This risk is minimized by keeping to well-known practice, but lowering the constraints on software development when available resources.
Table 6
Example quotations on agreement between designers and managers.

<table>
<thead>
<tr>
<th>Do It Differently</th>
<th>Keep This Technique</th>
</tr>
</thead>
<tbody>
<tr>
<td>&quot;You’re then down to, how do I ensure an adequate integrity of my product. Now then, I don’t see any reason why you wouldn’t want to strive for [high assurance]. Now how you do that is now up to you, because you’ve said [no requirement on assurance due to safety], but I still think we should strive for something approaching integrity [associated with the highest assurance] for almost everything that we do, because of the importance of it, functionality. So to take your example of using COTS software, if you’re selecting COTS software, there’s no longer any [Standard] requirements in play. If you think you have a good argument of why that software is good as is, and you don’t need to do anything else, then I think you make that case and then you say, I still think this is at this high integrity level.”</td>
<td>&quot;It’s tricky. You aren’t going to save much – if you took the [process] at [high assurance] and said, I’m going to use the same tools and techniques but I’m going to do [less assurance required], you’re hardly going to save anything.”</td>
</tr>
<tr>
<td>&quot;A good engineer is cautious and nervous and will not compromise … Because the problem with [high assurance] is everything has to be qualified, tools as well. So you can’t rely on automated picture to code generation, that’s what killed [Project]. But of course in [no requirement on assurance due to safety] you can automate everything.”</td>
<td>&quot;When we do [lower assurance] projects we typically just routinely just do [Practice A] anyway, because that’s how all – that’s how culture, that’s how engineers are brought up …”</td>
</tr>
<tr>
<td>&quot;So you know, [Practice A], particularly artefacts on the left hand side of the V, I think is a very good use of time and money.”</td>
<td>&quot;Well isn’t – I think my answer was, that I don’t think we – if it wasn’t mandated, I’d still do [Practice A].”</td>
</tr>
<tr>
<td>&quot;No, because I think we’d end up with [Practice A] anyway. Because I think if you needed it, I think we’d just do it anyway, as a matter of course, even if the process … We’re doing it because it’s useful ….”</td>
<td>[no requirement on assurance due to safety] I still think we should strive for something approaching integrity [associated with the highest assurance] for almost everything that we do, because of the importance of it, functionality. So to take your example of using COTS software, if you’re selecting COTS software, there’s no longer any [Standard] requirements in play. If you think you have a good argument of why that software is good as is, and you don’t need to do anything else, then I think you make that case and then you say, I still think this is at this high integrity level.”</td>
</tr>
</tbody>
</table>

research question by identifying a generalizable conflict within the community of software designers.

5.2. Characterising the conflict

In this subsection we address our second research question, characterising the nature of the conflict.

To explain the situation we note that tactical and strategic designers agree on the need for single-loop learning. As exemplified in Table 5 designers highlighted that the average engineer would not refer to standards frequently, with the risk of processes pigeonholing engineers or providing a false sense of security. In response, both groups work within the framework of guidance provided by standards to shape the practice of the average engineer at the firm in Table 7: tactical designers discuss the lack of an effect when dropping certain practice, while strategic designers highlight how this might mean that life-cycle costs are forgotten. There is also no disagreement that safety should guide this learning, as everyone interviewed at the firm agreed that this organisational value was paramount. The conflict is thus not a matter of organisational single-loop learning.

The conflict is arguably rather a matter of organisational double-loop learning in line with the external path to influence consensus outlined by the analytical framework. The norm in flux is whether to aim for aligning with the well-established practice or the most promising evolving practice in the inter-organisational community of practice. The former would already be explicitly coded into standards, while the latter would be allowed by (but not necessarily explicitly coded into) them.

This means that even if only designers and testers possess the technical expertise to interpret standards, software levels carry broader implications for safety management.

6. Discussion

This section addresses the third research question by discussing our analysis. It is thus our evaluation of the implications of our results for both practice and theory. In regard to practice, we contribute by discussing the implications of the identified conflict for the studied engineering context. In regard to theory on safety, there is a dearth of knowledge on engineers outside operational contexts, for instance regarding those who work in engineering organisations developing CPS products. Therefore, rather than building directly on the theoretical framework used for the analysis, we contribute to the theory on safety by linking our results to more generic concepts in the discourse on resilience.

6.1. Contribution to practice

This subsection starts by evaluating whether firms might be affected in different ways by the conflict described in the previous section. After identifying how a difference in organisational values should make it more likely that a firm supports the perspective of either tactical or strategic designers, we discuss the implications of this on safety management.

6.1.1. Early adopters and the majority of firms

As seen in Tables 6 and 7, the interviewees provided examples of how practice could – in their opinion – be changed for the better as the importance of lower software levels increase. It should be noted that this study does not focus on identifying the “correct” choices in this regard. Many of the suggestions were based in tacit knowledge, making a persuasive argument for adopting them outside the reach of this study. Regardless, such an aim would have presented at least two other major problems. Firstly, it could be in the interest of a firm to spread news about promising practice later rather than sooner, in case it provides a competitive advantage. Reporting on promising practice could thus be difficult, even if it was identified. Secondly, even if a practice is accepted by an increasing part of a community, it could be failing in ways that are not obvious to practitioners. The associated failures could be rare or difficult to detect, meaning that a significant amount of field data would have to be gathered to make an objective evaluation of a change to practice.

In fact, as the identified conflict is related to double-loop learning, whether practice is eliminated or transformed will most likely depend on the mission statement of a firm (Ard et al., 1997; Koryak, 2018). This statement articulates the firm’s organisational values, including the priorities it assigns to organisational states and outcomes such as innovation and predictability. Organisational values that emphasise innovation should give strategic designers a stronger position in the negotiation of practice within their community in the identified conflict, and vice versa. In fact, even if managers overcome the difficulties in acting independently associated with a lesser in-depth technical understanding of standards as pointed out in Table 5, they would still be likely to act in step with the community of designers: as they share these organisational values with designers, they are likely to assess arguments in a similar way. What our study indicates is thus that software levels will have different implications for practice in regard to the adopters that make early use of new technology to gain disruptive benefits and the majority of firms that do not make it their mission to lead on technology (Törngren et al., 2017; Geisberger and Broy, 2015). The former are more likely to embeden the strategic designers that seek to adopt novel practice. The latter are more likely eliminate practice when diversifying development between several software levels. In fact, the latter might even carry on without changing practice,
Table 7
Example quotations from strategic and tactical designer roles.

Tactical: Eliminating a Practice – Acceptable

“Tactical: Eliminating a Practice – Acceptable

We could develop [lower assurance] with the same type of [Practice A] that we use for [higher assurance]. My view is that we don’t need to do that and that therefore, when I think we are in some ways doing that already with the [CPS Product], we are.”

“I think we tend to apply it too early in a product development, certainly when it’s pre-release.”

“Oh no they wouldn’t, they wouldn’t, because they wouldn’t have time. There would be – even without knowing the details of the project, I can guarantee there would be time pressures and so if you told people they don’t have to do a particular part of the process, they will happily not do it.”

“I think there are probably some more that we could take out, it’s a matter of, you know, taking the time and just working out what’s actually required. There’s perhaps not enough benefit to spend the activity time doing that now, or it doesn’t, maybe there is, but it doesn’t seem obvious to do that.”

“Yeah, I think again, this is probably no surprise that I’ll say, yeah, that would be entirely acceptable. There are aspects of our [Practice B] process that even seem too rigorous for what we do for the [high assurance] stuff, so you know...

“...you would assume that’s fine, because we’ve taken the decision that even if we don’t find this even if this gets through to a later stage in the development process, so that it’s only picked up at verification, or possibly even after that, that’s okay, because we’ve decided that the [assurance required] of this software is such that that’s okay...”

“Some are based on experience of... sometimes you can kind of say, “Well, if I didn’t do this to any kind of great degree would it make any difference to the end product?” And you could probably say, “Well no, it wouldn’t.”... whereas other things, doing a part of a design or something, you might say, “Well, yes. That does have a big impact on that.” But it’s kind of down to experience I suppose and engineering judgement in a way.”

“In terms of the consequences, it depends what you’re looking, sort of the specifics or more general... specifics, it’s probably difficult to pin that down.”

“So you know that, even if we didn’t do the [Technique A], a greater proportion of things would get through, but some things get through anyway, so yeah, I don’t think you can ever, I don’t it’s easy to say there’s a direct link between one activity, but they all, overall add to the – so it’s an incremental thing isn’t it.”

Tactical: Task Risk Unknown

“A: But certainly, the [Low Assurance Development] team – I think it’s just something that we all know where we’ve got to get to and the team works together to get there... it’s very varied... people with 5 years’ experience; 10 years, 20 years. It’s a very, very diverse team... We have regular team meetings and issues like that will be kind of flagged up at the team meetings and probably discussed and then people will go away and kind of discuss those things outside of the meeting as appropriate.

B: So a lot of these issues are really solved by informal discussions?

A: Yeah.

“...So there’s different aspects to this aren’t there, so even if the process doesn’t mandate [Technique A] in general for a certain kind of activity, and even if you were an organisation that never did anything that was, you know, [high assurance development], with newer people you still, as part of how you train them and get them... I’ve worked with them... but that’s as part of their development within the organisation, it’s not because you’re following mandated process that says you have to do it... I’m just saying that because a process says you don’t have to [Technique A], that doesn’t mean that it wouldn’t still be part of training people and starting to deploy them onto projects and stuff like that.”

“So that’s still the senior members of the team going round and talking to more junior members. So it’s much more informal... It’s much more informally going and talking to people a couple of times a week and saying, “Hey, are you doing it? How are you getting on? What are you working on?” and having that discussion with people. If, through that discussion, you think that somebody’s maybe over-egging a solution, maybe doing something a little bit more complicated, over-considering what they’re doing, then you kind of bring them back down a bit and say, “Well actually, don’t worry about that. Concentrate on this.” Or equally, if somebody’s maybe not documenting everything or not doing what you think they should be doing according to the process, then kind of guide them at that point and say, “Well actually, you need to do more documentation. You need to document that particular aspect or do this sort of test. “...The coaching aspect is there, but it’s informal coaching a couple of times a week.”

Strategic: Keep Due to Quality and Life Cycle Costs

“Strategic: Keep Due to Quality and Life Cycle Costs

So, my interest is not so much in trying to, if you like, drop the level of software down to something that is more economic for producing bits of software, but it’s to try and raise the level at which we can do software in an economic way, because there’s a lot of evidence, and I think I put a reference to an article in my thing there, that if you do things rigorously, it may cost you more upfront but it will save you a lot on the maintenance cost, right, because you get fewer problems in the field, which are terribly expensive... So I think we should be very careful to make sure we look at life cycle costs rather than development costs. That’s the bit that worries me at the moment with all this discussion is we’re not looking at the whole life cycle; we’re just looking at the very immediate development cost.”

“Yes, the business requirements are still high, so you might adhere, you might choose to adhere to everything that is actually down there for [high assurance], even though you’re doing a [lower assurance] job, just because you want that level of confidence in the product...

“Because of course for them the impact of getting it wrong and getting it out into marketplace was – had massive business impact, just the same as it does in many safety critical ones. So the reasons for doing high integrity software are far broader than pure safety.”

“I think the [Practice A] gives a lot of benefit around the quality side of things, and finding the issues still earlier in the life cycle.”

“I suspect that’s one that you could lose very quickly... and I think in most cases it’s failed because actually when you do [Practice B] light what tends to happen is people stop doing [Practice B] and they do get into all the kind of messes that [Practice B] there to help you with. So I think people would give it up quite readily, but I think you’d have all the cost of not doing it properly.”

“Well the experience I’ve got of the [lower assurance] one it didn’t really save us much, and you lost that not doing [Practice A] or something, lost you something.”

“And there’s a lot of evidence that – I mean [Manager] again collected a lot of evidence in this area, saying that [Practice A] paid for itself.”

“I think that it would be very interesting for the business to reflect on the experience on [project] where they’ve been trying to run outside of [Practice B] for quite a long time. And it certainly – it obviously has its positives, but it certainly seems to have some significant negatives as well. So I think there’s got to be balance somewhere and I’m not sure we’ve got it.

“Well, I think the evidence is that if you don’t do some form of [Practice B] then you can lose track of what’s going on very rapidly.”

“I think we’ll, well there’s been several projects where the [restructuring project] cost hasn’t, I don’t think, hasn’t met what the business thought it was going to be...”

“Strategic: Cost Saving is not Direct

“It’s not quite as simple as that because actually an awful lot of the complexity and the difficulty of the higher levels is involvement with the regulatory authorities. So actually going to [lower assurance] say, obviously you don’t want to actually liaise with the regulatory authorities and that could take an awful lot of complexity and lead time out of the job.”

“My argument is we should be able to bring in alternate means as soon as we get down to [lower assurance] and below, particularly if we actually don’t want to develop the software ourselves, if we actually want to bring in COTS components or software developed elsewhere or software of uncertain pedigree, we don’t have process evidence for those. Typically we don’t have process evidence for software acquired via a means other than develop it ourselves, but it might be the most cost-effective way of getting that particular

(continued on next page)
product. So having the ability to have some sort of in-service reliability metric, say, for software and use that as a means of saying, "This software is good enough.""

"I think that what we haven’t touched on, I don’t think, is opportunities. So there are — we’ve talked about variation of [assurance] based upon process objectives and whether people can switch between them, etc., etc., etc. There are product features that are enabled by [lower assurance]. So there are things that — this is probably a debate actually. Does [high assurance] mean deterministic? So there are features that we’re talking about putting in our [product] that are things like algorithms that iterate to solutions and the more time you give them to iterate, they will get a more accurate solution, but less time it’ll be a less accurate solution but it might still be valuable, etc., etc. So there are soft real time and processor resource-based functions that we might be able to take advantage of ..."

"Yeah, I mean you know, you said that if we take something like the [COTS Component] that we’re not really striving for [high assurance] in that. I don’t see it like that, I think we are. I think when we take something like the [COTS Component], we’re actually saying, I think there’s a lot of, what we would’ve called in the [Standard] World Service History, associated with the [COTS Component], therefore if we use this particular version of it, we’ve got this evidence about how stable that is, and how good that is and what the bugs are in that."

"And, you know, that sometimes the potential savings of working at a [flow assurance] are exaggerated and the overhead of working to [high assurance] is not necessarily as much as people sort of feel it can be."

6.1.2. Splitting the community

Firms that choose not to lead on technology will see less of an impact at lower software levels and be less willing to move ahead of other firms in the market. They are thus likely to support the perspective of tactical designers. By being more likely to allow a differentiation between software levels, with the software at each level being substantially different in size and importance, it is also more likely that firms that do not lead on technology will physically split the community of software engineers between levels. As mentioned in Section 5.1, engineers with a background from outside the safety-critical industries could even be recruited to lower levels to avoid newcomers insisting on practice enforced at higher levels of assurance. It is unlikely that this isolation would be complete, as both communities would likely be coordinated by the same safety management. However, knowledge will no longer flow as easily between community members.

Engineers working at higher levels in the majority of firms will thus be less exposed to trends such as smart functionality. They will not need to make the associated judgement calls and thereby become less prepared for the day when the technological trend makes it into highly safety-critical functionality. As an example, the optimization of fuel consumption is a key selling argument in the aviation domain. The introduction of AI could be used to lower fuel consumption. One can also envision future functionality where this AI is used for adjacent reasons, such as how to handle fuel during an emergency landing. However, as long as the limitations of this optimization cannot be fully assured, an aeroplane will – to ensure safety – still have to be weighed down by carrying a non-optimized amount of fuel. If the gain is thus limited, the introduction of such functionality can be delayed. This would mean that engineers working at higher levels of assurance would not engage with associated strategic capabilities, such as the new practice and certification arguments needed to develop and evolve features based on AI.

To avoid a situation where the resilience of a firm is left vulnerable by engineers working at higher levels not knowing what to expect and look for, safety management should support the identification, organisation and training of informal opinion leaders working at different software levels. By being aware of differences across the organisation, gatekeepers (Paul and Whittam, 2010) and organisational liaisons (Tushman, 1977) can then eventually act as “knowledge transformers” (Whelan et al., 2009) able to help engineers absorb information from other levels.

6.1.3. Organising support across the industry

The process of standardisation in the safety-critical domain is not without issues — it can exclude relevant experts and has to date failed to make many underlying rationales explicit (Habli, 2017). The current shift towards more explicit safety argumentation in standards and by engineering organisations seems likely to succeed (Hawkins, 2013; Habli, 2017; Knight, 2014). In the future it is thus more likely that arguing for alternative means of compliance, as allowed by some standards, will come down to providing sufficient evidence of equivalence between the novel and replaced practice in regard to an explicitly defined goal. The challenge for firms will be to ensure that the evidence is plausibly complete, i.e. that all corner cases have been considered. While current standardisation processes itemize applicable techniques well by utilizing a few experts (Knight, 2014), the challenge of identifying corner cases is more suitably solved by including a broad set of participants. The identified conflict centred on software levels makes this involvement less likely to happen, as a large set of the relevant CoP will rather have to focus on meeting the needs of smart functionality by eliminating practice. This would involve a significant effort in evolving the system safety assessment processes that determine software levels and the product architectures that enable their diversification. This will make it more difficult for early adopters of technology to get the feedback they need to ensure they stay resilient when introducing novel practice.

Another concept to CoP that relates to group activity in organisations is crowdsourcing, which describes the use of information technology to elicit crowds to address an intra- or inter-organisational open call (Zuchowski, 2016). A difference to CoP is that expertise is less emphasised, often meaning that pre-selection or limitation of participants is viewed with scepticism (Zuchowski, 2016; Geiger, 2011; Simula and Ahola, 2014). Originally crowdsourcing was used to tap into the “wisdom of the crowd” (Surowiecki, 2004), but crowds have since then also been used for expert problem-solving (Feller, 2012). This activity includes several phases of which the most known is probably the evaluation and aggregation of results, which frequently involves crowd voting (Zuchowski, 2016). A known risk with crowdsourcing is thus the pressure to conform to the opinions of peers (Stieger, 2012), which may lead to certain users dictating the responses of crowd members. However, crowdsourcing also contains phases where the focus is on developing ideas in groups, and the evaluation and aggregation can be performed by select experts. As a complement, crowds could in this way offer a possibility to analyse and discuss the strengths and weaknesses of novel practice across a larger group. Safety management at a firm that is aiming to align with promising evolving practice in inter-organisational CoP would thus do well to establish at least an internal crowdsourcing capability for cooperation on ideation analysis and refinement.

In a similar way, additional benefits could be achieved by extending the crowdsourcing to other firms with the intent to support standardization. Firstly, this would allow certifiers to learn from the discussion on assuring novel practice earlier, lowering the risk of using novel assurance argumentation. Secondly, it would provide more examples of how to handle novel practice in the standardization process, clarifying the interpretation of standards. Thirdly, it would make the standardization process less reliant on individual experts, especially if digital
tools for using the power of large crowds were adopted. Fourthly, if large firms involved many of their employees, it would arguably both pressure and make it possible for standardization organisations to lower the cost of gaining access to standards. The largest issue with such an approach would arguably be how to deal with intellectual property with so many employees from different firms involved (as this is a common problem when large firm-external crowds are used (Zuchowski, 2016; Schenk et al., 2017). However, as safety argumentation to some extent relies on openness to be accepted by certifiers, there is an incentive for firms to allow some risk in this regard. Furthermore, ensuring this openness would also be important to avoid the negative effects alluded to in the previous paragraph, such as peer pressure, rumour spread or use of made-up numbers. Keeping the evaluation of the crowdsourcing separate from the crowd would allow experts to evaluate whether the supplied arguments and evidence are objective or not.

6.2. Contribution to theory

The perspectives of the two different groups of designers are well aligned with two models on how rules primarily support safety that have been discussed for operational contexts in the discourse on resilience. According to Model 1 rules “define and guide behaviour in complex and often conflicting environments and processes” (Hale and Borys, 2013), while Model 2 sees them as “supports, not strait jackets; as tools to coordinate and structure creativity and innovation, not as controls to limit freedom” (Hale and Borys, 2013). As seen in Table 7, strategic designers gravitated towards a Model 1 perspective, arguing that there were many bad examples from breaking the “rules” of the standards. The tactical designers instead adopted a Model 2 perspective, arguing that expert engineering judgement could usually lead to acceptable interpretations of these “rules”.

While there is a wealth of research on these models in operational contexts, little is known about how they relate to standardization (Hale and Borys, 2013). We suggest that the results of this study are helpful in this regard. Perspectives adopting Model 2 are often found where firms have to rely on the ability of operators to steer their organisation away from unsafe situations (Hale and Borys, 2013; Rasmussen, 1997; Dekker, 2006). These firms come to depend on collegial patterns of authority and well-informed decentralised decision making that allow operators to act autonomously (La Porte, 1996). In complex domains guidance is thus likely to be more usable if specified in process rules rather than action rules (Hale and Borys, 2013). Arguably, the expectation could be that causality would also work in the opposite direction – that firms operating in a complex development environment would align more with Model 2 if more leeway was given to engineering judgement, for instance by less prescriptive standards. Firms could be expected to develop their resilience (Hollnagel, 2006) by enabling engineers to individually overcome the problems exemplified in Table 5, i.e. the risk of processes pigeonholing or providing them with a false sense of security. In other words, that the individual engineer would be tasked with more responsibility for the organisation’s ability to ensure an understanding of what to expect, what to look for and what to do to ensure safety as circumstances change.

However, our results show that rulemaking does not automatically shift towards Model 2 in these circumstances. The perspective on rulemaking that an engineering firm adopts is more likely to be related to whether it is an early adopter of technology. Only then do standards become important by either allowing novel practice to be imported or not. Less prescriptive standards might thus mean that a firm puts less, rather than more, stock in their engineers’ ability to decide autonomously what to do to ensure safety.

7. Conclusions

Our results suggest that there is a potential for conflict when the consensus on practice for software engineers used to primarily working at higher levels of assurance is eroded by trends that increase the importance of lower software levels. This conflict materialises between tactical and strategic software designers, as the different risk foci of these groups make them advocate elimination and transformation of practice, respectively. The nature of this conflict can best be described as a matter of organisational double-loop learning, where the norm in flux is whether to aim for aligning with well-established practice or the most promising evolving practice.

As the conflict is centred on double-loop learning, it will have different implications for adopters that make early use of new technology to gain disruptive benefits than for the majority of firms that do not make it their mission to lead on technology. For the majority of the firms the risk is that engineers working at higher levels are caught unprepared. To avoid a loss of resilience, safety management should support the identification, organisation and training of informal opinion leaders working at different software levels. For early adopters the risk is that they do not get the feedback they need to ensure they stay resilient when undergoing change. To compensate, safety management should explore crowdsourcing for analysing and discussing the strengths and weaknesses of novel practice. If crowdsourcing for this purpose could be extended across entire industries, it should provide benefits to the entire engineering standardization process.

The contribution to theory by this study indicates that less prescriptive standardization does not automatically shift rulemaking towards allowing operators to act more autonomously. Organisational values, such as the willingness to lead on technology, are more important in establishing the perspective on rulemaking. Only after these values have established a perspective on rulemaking will standards become important by either allowing novel practice to be imported or not. Less prescriptive standards might thus mean that a firm puts less, rather than more, stock in their engineers’ ability to decide autonomously what to do to ensure safety.

Declaration of Competing Interest

None.

Acknowledgements

The research leading to these results has received funding from Sweden’s Innovation Agency (VINNOVA) under grant agreement n° 2015-01524. Special thanks go to Vicki Derbyshire for her help in proofreading.

References


Brinkmann, S., Kvale, S., 2015. The Social Construction of Validity, in InterViews,