Safety Practice and its Practitioners: Exploring a Diverse Profession

Mario Gleirscher\textsuperscript{a,c}, Anne Nyokabi\textsuperscript{b,c}

\textsuperscript{a}Department of Computer Science, University of York, United Kingdom
\textsuperscript{b}Process Industries and Drives Division, Siemens AG, Germany
\textsuperscript{c}Formerly: Institut für Informatik, Technische Universität München, Germany

Abstract

Context: System safety refers to a diverse engineering discipline assessing and improving various aspects of safety in socio-technical systems and their software-intensive sub-systems. While system safety has been a vital area of applied research for many decades, its practice and practitioners seem empirically still not well studied. Beyond anecdotal evidence—case reports, interviews, discussion forums, blogs, and “war stories” serving as interesting examples in textbooks—and surveys, we are missing open, large-scale, and long-term empirical investigations that promote knowledge transfer and research validation.

Objective: In this article, we explore means for work that safety practitioners rely on, factors influencing the performance of these professionals, and their perception of their role in the system life cycle. Along with that we want to examine observations from previous research.

Methods: We build a construct of safety practice, collect data for this construct using an on-line survey, summarise and interpret the collected data, and investigate several hypotheses based on the previous observations.

Results: We analyse and present the responses of 124 practitioners in safety-critical system and software projects. Aside from other findings, our data:

• suggests that safety decision making mainly depends on expert opinion and project memory,
• lacks evidence that safety is typically a cost-benefit question,
• does not exhibit the prejudice that formal methods are not beneficial,
• leaves it unclear as to whether or not standards and methods have become inadequate,
• and
• indicates that safety is not typically confused with reliability.

Additionally, we contribute a research design directing towards explanatory empirical studies of safety practice.

Conclusions: We believe that empirical research of safety practice is still in an early stage, bearing the risk of undesirable mismatches of the state of the art and the state of practice. However, this situation offers great opportunities for research.

Keywords: Safety-critical system, safety practitioner, professional profile and situation, state of the practice, on-line survey, exploratory study

1. Introduction

System safety practice (safety practice for short) is a remarkably diverse field spanning many disciplines involved in the system life cycle, influenced by heterogeneous criticality-driven safety cultures\textsuperscript{[1, 2, 3]} across various application domains, geographical regions, and regulatory authorities.

Researchers have surveyed and investigated practised approaches to accident prevention, for example, in the chemical plant and nuclear power plant sectors\textsuperscript{[2]} and in the construction industries\textsuperscript{[3]}. However, our literature search has not uncovered a single officially published empirical investigation (i.e., a case or field study, a controlled field experiment, a survey of practitioners) of the effectiveness of practised approaches to prevent or reduce software and (control) systems’ contributions to hazards.

In the following, we highlight the motivations for our study, describe observations from previous research, outline our research objective, and summarise the contributions of this work.

1.1. Problem Statement

From exploratory content analysis of more than 200 selected question and answer posts on several safety practitioners’ (SP) on-line channels of a period of 4 years and one expert interview\textsuperscript{[4]}, we observe that SPs:

1. discuss various issues with the application of standards, calculation of failure rates, correct planning of safety tests, and completeness of hazard analyses;
2. are missing a standardised way of integrating safety with security activities;
3. are concerned about the adequacy of methods, a lack of safety education, and the misunderstanding of their role.
From exploratory content analysis of more than 370 case reports (i.e., on incidents and accidents) from the aviation, automotive, and railway domains and 7 semi-structured interviews with SPs from these domains [5], we observe that

1. human errors and specification errors were more often reported as accident root causes than software implementation errors—this is consistent with the findings in [6, p. 30f];
2. no IT security problems were reported;
3. reports in general, and comparably often in the automotive domain, were non-informative of subtle accident root causes (i.e., causes lying outside the possibilities, budgets, or obligations of accident analysts and investigators);
4. few of the selected reports at least suggest that accidental complexity [7]—particularly, missing or mistaken maintenance, refactoring, evolution, or migration—negatively affects system safety;
5. interviewees report issues of unclear separation of system-level and software-level activities (cf. [8]);
6. interviewees state that available methods are currently just appropriate in their domains but can easily get insufficient for complex future applications.

These observations fuel some almost negligently accepted computer-related risks—as regularly archived by Neumann et al. [9]—as well as occasional but recent worries about the state of the practice and education in safety engineering in particular [10] and in software engineering (SE) in general [11].

1.2. Research Objectives

These findings certainly ask for more evidence. In line with the research agenda in [12], safety engineering research might, hence, pose clarifying questions such as:

1. Which means are SPs familiar with and which do they currently use? How clear, unambiguous, consistent, up-to-date, and effective are those means?
2. What are the SPs' current problems, challenges, needs, and expectations?
3. How do SPs view their profession, role, and contribution in the life cycle?

1.3. Contributions

This work contributes to safety research in several ways:

• First, we present results of a cross-sectional self-administered on-line survey among SPs: Particularly, we sample some of their experiences, views, opinions, and their self-perception.

• Then, we test several comparative hypotheses (Section 4.4) on safety practice and SPs and interpret our test results (Section 5.1) with respect to findings and experience from previous work [5, 4]. This way, we also elaborate on results in [13].

• Furthermore, we respond to the request from Alexander et al. [14] and Rae et al. [15] for applying improved methodology in empirical research of safety practice, as well as the desire of a stronger involvement of SPs in research evaluation such as stated by Martins and Gorschek [12].

• Moreover, we contribute a research design (Section 3) for similar empirical assessments. This setting might as well be applicable to other SE domains (see, e.g. [16]).

1.4. Overview

Figure 1 provides an overview of the research method for this article. After discussing terminology and related work in Section 2 and describing our research method in Section 3, we present our results in Section 4. Particularly, we describe our sample in Section 4.2 and summarise the results of all valid responses in Section 4.3. Section 4.4 highlights the results of several hypotheses tests. Our discussion follows in Section 5, with the interpretation of our test results in Section 5.1 and the examination of threats to the validity of our study in Section 5.4. We summarise our findings in Section 6. Appendix A contains a detailed summary of the response data.
2. Background

We introduce important terms as well as related work we will revisit in our discussions below.

2.1. Terminology and Definitions

The life cycle of an engineered system typically refers to the phases of design, implementation, release, maintenance, operation, and disposal. Dependability then encompasses the handling of reliability, availability, maintainability, and safety in the life cycle, for example, by improving fault-tolerance.

In this work, we focus the discipline of system safety,1 including functional safety. System safety is usually situated in the context of safety of machinery,2 process safety,3 structural safety,4 or occupational health and safety. These disciplines have in common the identification, assessment, and management of operational risk. This procedure includes the prevention or handling of undesired events at any stage (e.g. hazards or safety risks, incidents, and accidents) and of any type (e.g. human error, software faults, and system failures). In addition, security of information technology (IT security or security for short) is the discipline of protecting computer-based systems and data against malicious attacks and unauthorised access.

Then, safety practice denotes the practical aspects of system safety in both industrial settings and applied research. Based on this, we consider a safety practitioner as a person who supports or performs safety decision making, particularly, by identifying hazards and assessing their causes and consequences, the design of hazard countermeasures (also known as hazard controls), the assurance of safety, or by performing research and consultancy for these safety activities. Importantly, there are many means—that is, best practices, methods, techniques, and standards—to apply in these activities.

2.2. Related Work

As indicated in Section 1, there are only few cross-disciplinary exploratory inquiries of safety practice and its practitioners. The following studies demonstrate the importance of empirical methods (interviews and related survey methods such as focus groups and questionnaires) in further examining safety practice.

Dwyer [17], for multiple disciplines, and Knight [8], for software engineering, characterise safety practice from their experience, forecasting the ongoing trend of increased automation, the increasingly critical interplay between the involved engineering domains, and the corresponding challenges for future safety research.

Adequacy of Means of Work in Safety Practice. Safety-critical systems are subjected to automation (i.e., the use of qualified and verified tool chains) for their development, testing, and overall assurance. Graaf et al. [18] and Kasurinen et al. [19] investigate challenges and obstacles to adoption of new methods, languages, and tools in embedded system RE, architecture design, and software testing. Our study explores this direction within safety practice.

Hatcliff et al. [20] summarise particular challenges in the certification of software-dependent systems and suggest improvements, stressing the concept of “designed-in safety/security.” These works inspired and underpin our hypotheses but are different from our survey approach to examining safety practice and its practitioners.

Chen et al. [21] report on the challenges and best practices of using assurance cases. Our questionnaire about safety practice includes more general questions about methods, training, and interaction, backed by a larger number of data points.

For organisations that engineer safety-critical systems, Ceccarelli and Silva [22] provide a framework for compliance checking during and after the introduction of new safety standards (e.g. DO-178B) into an organisation. In our study, we are asking SPs whether safety standards known and used by our respondents, actually improve the organisation’s safety practice.

McDermid and Rae [10] report on their cross-domain insights into the practice of engineering safety-critical systems, discussing the question: “How did systems get so safe despite inadequate and inadequately applied techniques?” Without presuming that modern systems are acceptably safe, we interrogate SPs about their means of work.

Wang and Wagner [23] investigate decision processes in safety activities. For complex and highly critical systems such processes are usually committee- or group-driven to reduce organisational single points of failure. The authors examine whether such decision making is prone to a number of pitfalls known as “groupthink” and studied in group psychology. Being more exploratory in nature, our study design differs from the psychology-based construct used in [23] and yet addresses a fraction of it.

Process Factors Influencing Safety Practice. Requirements engineering (RE) and, particularly, requirements specification, are critical points of failure in every safety-critical system project. Examining research on the communication and validation of safety requirements in industrial projects, Martins and Gorschek [12] conclude that there is a lack of evidence for the usefulness and usability of recent safety research. We want to contrast their finding with how practitioners currently perceive the adequacy of their means of work.

Nair et al. [24] present results from a survey of 52 SPs on how they manage the variety of safety evidence for critical computer-based systems. Good evidence management
implies many technical challenges in safety practice. Particularly, traceability is crucial for change impact analysis (CIA), that is, the analysis of how changes of safety-critical artefacts (e.g. specifications, issue databases, designs) are propagated and whether these changes have negative safety impact. Borg et al. [25] report on 14 interviews with SPs about their CIA activities, finding that SPs have difficulties in understanding the motivation of CIA, are overwhelmed by the information they have to process when conducting CIA, and struggle with trusting and updating former CIAs. From a cross-sectional survey of 97 practitioners, De la Vara et al. [26] observe insufficient CIA tool support. Our study examines such means of work from a more general viewpoint.

In the Sections 3.3.1 and 3.4.1, we establish relations between these works and our study. In Table 8 in Section 5.2, we compare their findings with our results.

3. Survey Planning

This section describes our survey design (Section 3.1), the survey instrument (Section 3.3), our working hypotheses (Section 3.4), the procedure for data collection (Section 3.5) and analysis (Section 3.6), and instrument evaluation (Section 3.7). We follow the guidelines in [27] for planning and conducting our survey and [28, 29, 30] for the reporting.

3.1. Research Goal and Questions

From our project experience summarised in Section 1.1 and previous research [13, 5, 4], we have learned about typical issues in safety practice. This cross-sectional survey design aims at resolving these issues. The objective of this exploratory study is to investigate safety practice and its practitioners and to examine observations we made during our preliminary research.

For this, we explore three research questions:

RQ1 Which means do SPs typically rely on in their activities? How helpful are those means to them?

RQ2 Which typical process factors have influence on SPs’ decisions and performance?

RQ3 How do SPs perceive and understand their role in the process or life cycle?

3.1.1. Construct

For this objective and these research questions, we introduce the construct safety practice and its practitioners (SPP). This construct incorporates SPs’ processes, tasks, roles, methods, tools, and infrastructures and, by interrogating them via a questionnaire, their views and opinions of safety practice. SPP is divided into three sub-constructs: Classification of SPs, Constituents of safety practice, and

| Classification of SPs, Constituents of safety practice, and |
| Classification of SPs, Constituents of safety practice, and |

Table 1: Classification criteria for characterising the population and for sample assessment. Legend: MC...multiple-choice, (N)ominal or (O)rdinal scale.

| Classification Criterion | Scale |
|-------------------------|-------|
| Educational Background  | N / MC|
| Application Domains     | N / MC|
| Level of Experience     | O / duration in years|
| Familiarity with Standards | N / MC|
| Familiarity with Methods | N / MC|
| Geographical Regions    | Open / MC|
| Native Languages         | N / MC|
| Working Languages        | N / MC|
| Safety-related Roles     | N / MC|

Table 2: Constituents of safety practice and practitioner’s expectations and challenges. Legend: (N)ominal or (O)rdinal scale, (T)ruth values as nominal scale, ∗...half-open or open.

| Construct of Safety Practice | Scales |
|-----------------------------|--------|
| Safety Process (activities, roles, and practitioners) | N / e.g. decisions, hazard identification, resources |
| Factors (constraints and issues) | T / e.g. lack of resources, high schedule pressure |
| Means (conventional techniques; formal methods; tools; norms; skills; knowledge sources) | N* / e.g. FMEA, ISO26262, FMEA expertise, expert opinions |
| Application domains (current, new, complex) | N* / e.g. systems based on adaptive control, machine learning |

Expectations & Challenges in Safety Practice (as perceived by SPs)

| Expectation | Scale |
|-------------|-------|
| Performance of safety activities | O / high...low performance |
| Adequacy of means | O / high...low adequacy |
| Collaboration between safety and security engineers | O / effective...ineffective |
| Value of knowledge sources to SPs | O / high...low, per class of methods or standards |
| Adaptation and improvement of SPs’ skills | O / high...low |
| Notion, perception, and priority of safety activities | N* |
| Contribution of SPs to system life cycle | O / high...low contribution |

Expectations & challenges in safety practice. The construct is visualised in Figure 2.

The criteria for classification in Table 1 and the breakdown in Table 2 are results from the first author’s experience from research in system safety, from collaborations with industry, from expert interview transcripts, and from the supervision of the three thesis projects documented in [13, 5, 4]. The bottom-up creation of the SPP construct took place along the lines of grounded theory [31] based on these materials and further experience gained during the survey execution.

Below, we use the following prefixes when referencing important content items: RQ for research questions, h for working hypotheses, q for questions in the questionnaire, and F for findings. References will have the shape ⟨X⟩⟨Label⟩[−⟨o⟩] where X ∈ {RQ, h, q, F} and o can refer
to an answer option in the questionnaire. Additionally, we provide legends along with the corresponding figures.

3.2. Survey Participants and Population

Safety practitioners are our direct study subjects, our target group. A safety practitioner is a person whose professional activities as a practitioner or researcher in industry or academia are tightly related to the engineering of safety-critical systems. Table 1 lists criteria we use to characterise and identify members of the population of SPs. Safety practice, as described in Section 2.1, is our indirect study object. SPs participating in our study are also called study or survey participants or respondents.

3.3. Data Collection Instrument: On-line Questionnaire

Table 4 provides details on the (q)uestions we discuss in this work. For sake of conciseness, concept traceability, and compact presentation in this article, we consolidated the questions stated in our questionnaire, of course taking care of maintaining their original meanings. For verification of this transformation, the whole original questionnaire and its code book are documented in [13].

3.3.1. Motivations underlying the Questions

In the following, we establish links between the questions and other research summarised in Section 2.2.

qAdeqOfMthStd. Bloomfield and Bishop [32] contrast prescriptive regulation with goal-based regulation, reviewing current practice, highlighting potential benefits of safety cases along with the challenge of gaining sufficient confidence. Starting from a general position, question qAdeqOfMthStd is about norms adequacy in general.

qValueOfKnow. Lethbridge et al. [33] observe that test and quality documentation is the most likely maintained kind of documentation. With question qValueOfKnow, we want to find out about how project documentation is used in safety decision making.

Moreover, Rae and Alexander [34] examine how confidence in safety expert judgements (e.g. individual versus group judgements) is justified and leads to actual validity of the conclusions the further stages of the safety life

| Type          | Values               |
|---------------|----------------------|
| value         | very high (vh), (h)igh, (m)edium, (l)ow, very low (vl) |
| agreement     | strongly agree (sa), (a)gree, neither agree nor disagree (nand), (d)isagree, strongly disagree (sd) |
| impact        | (h)igh, (m)edium, (l)ow, (n)o impact |
| adequacy      | very adequate (va), (a)dequate, slightly adequate (sa), not adequate (na) |
| frequency     | often, rarely/occasionally, never; or all, many, few, none |
| choice        | single/multiple: (c)hecked, (u)nchecked; or yes, no |

For maturity measurement, Ceccarelli and Silva [22] work with a construct similar to the one in Table 2. By asking question qAdeqOfMthStd, we cover practitioners’ views (and opinions) independent of a specific norm.

The questions about adequacy of means (particularly, qAdeqOfMthStd, qAppOfMth, qPosImpOfFMs), aim at the re-examination of known challenges as, for example, discussed by Kasurinen et al. [19] and Graaf et al. [18]. The answer categories for question qAdeqOfMthStd are based on industry sectors with a relatively high pace of innovation and/or new, complex, but not yet well-understood system applications (e.g. self-driving cars).
cycle are based on. The authors argue that expert risk assessments exhibit low effectiveness in measuring risk as an objective quantity and propose “risk assessment as a means of describing, rather than quantifying risk.” Their analysis extends the background of qValueOfKnow.

**qIntOfSafSec and qPrioOfSafety.** While Chen et al. [21] focus on the aspect of training and collaboration in safety assurance, our study crosses these aspects generally with the questions qIntOfSafSec and qPrioOfSafety about interaction in and efficiency of safety activities.

The questions qIntOfSafSec, qValOfContrib, and qCoWorkers address the integration of safety activities with the life cycle, similar to Bjarnason et al. [35] on the alignment of RE and verification and validation.

In contrast to tool support for optimal auditing as investigated by Dodd and Habli [36], our questions (i.e., qEffRoleOnJob, qValOfContrib, qCoWorkers, and qImpOfExp) help to solicit personal views of SPs as external auditors and consultants.

**qImpOfConstr, qImpOfEco, and qNotionOfSafety.** As summarised in Section 1.1 and as discussed in [5], we presume negative consequences of “accidental complexity” [7] on system safety. Lim et al. [37] examine the perception of technical debt, highlighting the inevitable trade-off between software quality and business value. In an unfortunate case, an acceptance of technical debt can lead to an acceptance of low software quality, and for some systems, to an acceptance of accidental complexity. Whenever this reasoning applies to a safety-critical system, we should ask whether this system is taken in by an unacceptable trade-off between safety and business value? Asking the questions qImpOfConstr, qImpOfEco, and qNotionOfSafety, we inversely probe the demand for investigations of the safety impact of technical debt.

Based on the SPP construct, we interrogate SPs about supportive factors (qValueOfKnow, qPrioOfSafety) and obstacles (qImpOfConstr) in safety decision making, gathered from our previous interviews in [5, 4].

### 3.3.2. Notes on the Questionnaire

Some questions in Table 4 are half-open, that is, we allow respondents to extend the list of given answer options. The scales used for encoding the answers in the column “Scale” are described in Table 3. We treat value and agreement as a 5-level LKERT-type scale. Value, impact, adequacy, and frequency scales are equipped with a “do not know (dnk)” option. Together with “neither agree nor disagree (nand)” answers, participants are given two ways to stay indecisive. This way, we try to reduce bias by forced responses. From comparative analysis, we conclude that it is safe to discard dnk-answers and missing answers from our analyses.

We expect survey participants to spend 20–30 minutes
on the questionnaire. Although we do not collect personal data, they can leave us their email address if they want to receive our results.

3.4. Working Hypotheses

We derive working hypotheses from our observations (Section 1.1) from previous research [5, 4, 13]. Table 5 contains two types of working (h)ypotheses we want to analyse and test with the data we collect from the survey participants. First, the base hypotheses incorporate observations, assumptions, or prejudices, either identified from our previous research or already made by other researchers. Additionally, we elaborate comparative hypotheses during exploratory analysis [38] of the responses.

Some hypotheses in Table 5 are directly measured by a single compound question (see, e.g. hExp:DivGTSing and qImpOfExp). We do not collect data for each individual construct referred to in such hypothesis-question pairs.

Figure 2 summarises the survey design presented in Sections 3.1 to 3.4 by showing important interrelationships between the base hypotheses, the questions of the questionnaire, and the parts of the SPP construct.

3.4.1. Motivations underlying the Hypotheses

In the following, we justify our working hypotheses through establishing links to other research (Section 2.2).

hExpDecides: SPs’ activities mainly depend on expert opinion and experience from similar projects. It is well-known that experts are fallible (see, e.g. recent investigations in [39, 34, 23]) and, thus, relying on experts in organisational (and engineering) decision making can contribute to critical single points of failures in such organisations. Moreover, it is well-known that reusing (e.g. cloning) repositories from finished projects in similar new projects bears many risks of errors in reuse or update of these data. Our previous interviews suggest that both these knowledge sources are used in safety practice.

hLoResLoSaf: A lack of resources has a negative impact on the performance of safety activities. The observations in Section 1.1 motivate the collection of evidence on whether the performance of safety activities.

hInsufStds: Conventional methods (e.g. FMEA, FTA) are challenging to apply to complex modern applications. The observation that conventional methods have become inadequate is broached by Knight [8, 42]. Likewise, McDermid and Rae [10] and Hatcliff et al. [20] underpin hInsufStds and hInsufMeth, though not the long-standing [43] and frequent expectation that formal methods (FM) have a positive impact on safety practice (hFMsImpSaf).

hFMsImpSaf: The use of formal methods has a positive impact on the performance of safety activities. The efficacy of FMs in practice has been an only moderately researched subject for many years, investigated, for example, by Barroca and McDermid [44] and Woodward et al. [45]. One intention underlying hFMsImpSaf is to determine whether we have to further examine FM effectiveness to cross-validate reported experiences (e.g. [46]).

hSafBySec: For current applications, the assurance of safety also depends strongly on the assurance of security. Safety-critical applications of networked or connected (software) systems have recently revived the question of how safety and IT security influence each other? Along these lines, the justification of hSafBySec is based on manifold anecdotal evidence (see, e.g. [47]) that security problems can cause safety violations and, possibly, vice versa.

hSafIsCost: Safety is more seen as a cost-increasing rather than a cost-saving part in many application domains. How are the practical achievements and implications of system safety and the effort spent therefor related? How relevant are such utilitarian and controversial questions to SPs and their organisations? Touching this subject, hSafIsCost is formulated in the context of “total cost of safety,” that is, the cost of accident prevention and accident consequences borne by organisations that engineer and operate safety-critical systems. hSafIsCost’s truth might contribute negatively to the role of SPs in an (engineering) organisation.
Table 5: Overview of hypotheses (used as $H_i$ in the tests). **Legend:** See Section 3.3. The quantification ranges a–j refer to the answer options of the questions associated with the hypotheses, see Figures 12 to 23. The original questionnaire is documented in more detail in [13].

| Hypothesis | Supported if… (AC, Section 3.6.2) |
|------------|-----------------------------------|
| **Base Hypotheses** | |
| hExpDecides: SPs’ activities mainly depend on (d) expert opinion and (g) experience from similar projects. | $\forall o \in \{d, g\}: med(q ValueOfKnow\^{o}) \in \{h, v\h\} \land o \text{ among } 3 \text{ highest values (of 7) knowledge sources } \wedge med(q ImpOfExp\^{o}) \in \{a, sa\}$ |
| hLoResLoSa: There is a lack of resources that has a negative impact on the performance of safety activities. | $\forall o \in \{d, a\}: med(q ImpOfConstr\^{o}) \in \{m, h\} \land med(q ImpOfEco\^{o}) \in \{often\} \land q NotionOfSaf\^{o} \leq 30$ |
| hInsuStdSaf: Safety activities for highly-automated applications lack support of appropriate standards and methods. | For $\geq 5$ out of 7 domains $o$: $med(q AdeptOfMthStd\^{o}) \in \{sa, na\}$ |
| hFmsImprSaf: The use of formal methods has a positive impact on the performance of safety activities. | $med(q AppOfMeth) \in \{a, sa\} \land q PosImpOfFMs_{m+h} > 25\%$ |
| hSpsAdptSks: SPs improve their skills towards new applications, e.g. by studying recent results in safety research. | $\forall o \in \{a, b\}: med(q ImpOfSkills\^{o}) \in \{a, sa\}$ |
| hSafBySec: For current applications, the assurance of safety also depends strongly on the assurance of security. | $\forall o \in \{a, c, e, f\}: med(q IntOfSafSec\^{o}) \in \{a, sa\}$ |
| hSafsCost: Safety is more seen as a cost-increasing rather than a cost-saving saving in many application domains. | $q NotionOfSaf\^{a} > 60\% \land \forall o \in \{b, e\}: q NotionOfSaf\^{o} < 40\%$ |
| hLoCollLoSa: A lack of collaboration of safety and security engineers has a negative impact on safety activities. | $\forall o \in \{h, i, j\}: med(q IntOfSafSec\^{o}) \in \{a, sa\}$ |
| hHiPrioHiSaf: Prioritisation of safety in management decisions enables SPs to perform their tasks more efficiently. | $\forall o \in \{a, b\}: med(q PriorOfSafety\^{o}) \in \{a, sa\}$ |
| hSafsRel: SPs understand safety as a special case of reliability. | $\forall o \in \{a, e\}: med(q UndesredEd\^{o}) \in \{a, sa\} \land \forall o \in \{b, c\}: med(q UndesredEd\^{o}) \in \{d, sd\}$ |
| hSafsValued: SPs believe that their non-safety co-workers attribute high value to SPs’ contributions. | $med(q ValOfContr\^{a}) \in \{h, v\h\} \land med(q ValOfContr\^{o}) \in \{m, h\}$ |
| hPosSelfimg: SPs perceive their contribution as highly valuable. | $h$ |
| **Comparative Hypotheses** | |
| hExp:DivGTSing: SPs with high diverse expertise better perform in safety activities than SPs with low singular expertise. | One-sided $U$ succeeds with $p < 0.05$ |
| hVal:SenLJUn: Senior SPs attribute lower value to their role in the system life-cycle than junior SPs (cf. hSafsValued, hPosSelfimg). | One-sided $U$ succeeds with $p < 0.05$ |
| hAdapt:SenGTJun: Senior SPs agree more than junior SPs that skill adaptation (e.g. learning) is required and takes place (cf. hSpsAdptSks). | One-sided $U$ succeeds with $p < 0.05$ |
| hAdapt:AutoGTAero: SPs using automotive standards agree more than SPs using aerospace standards that skill adaptation (e.g. learning) is required and takes place (cf. hSpsAdptSks). | Two-sided $U$ succeeds with $p < 0.05$ |
| hInsuMeth:EngDiSci: Engineering-focused SPs agree different from research-focused SPs with hInsuMeth. | One-sided $U$ succeeds with $p < 0.05$ |
| hInsuMeth:AutoGTAero: SPs using automotive standards agree more than SPs using aerospace standards with hInsuMeth. | |

**hLoCollLoSa:** A lack of collaboration of safety and security engineers has a negative impact on safety activities. According to Conway [48], the structure of an engineered system converges towards the (communication) structure of its engineering organisation. For example, in a safety-critical distributed embedded system (e.g. avionics, process automation, and automotive architectures), teamwork collaboration would determine the architectural decomposition and direct communication links in the architecture. However, team collaboration not necessarily implies keeping track of the impact of critical changes across all critical relationships. It is also known that critical relationships in a complex architecture are far from obvious. Sadly, such relationships are sometimes only indirectly perceived as an undesired emergent property. Hence, we ask SPs about the collaborations between so-called “property engineers,” e.g. safety and security engineers (qIntOfSafSec).

**hSafsRel:** SPs understand safety as a special case of reliability. Leveson [49] stresses an observed misconception about system safety, namely that the responsibility to make systems safe enough is reduced to the responsibility to make their critical parts just reliable enough. Her claim stimulates the question to which extent SPs are solely driven by reliability concerns and which negative implications this might have. Moreover, hSafsRel is also motivated by examinations [50] of how findings from previous accidents can be included in safety arguments.

**hAdapt:AutoGTAero:** SPs using automotive standards agree more than SPs using aerospace standards that skill adaptation is required and takes place. During our interviews we heard several times that system safety practice in the automotive domain is for several reasons less developed than in other domains, such as aerospace. Hence, we assume...
that automotive SPs are currently more strongly involved in or aware of skill development in their domain than SPs in aerospace.

3.5. Data Collection Procedure: Sampling

To draw a diverse sample of safety practitioners, we
1. advertise our survey on safety-related on-line discussion channels,
2. invite practitioners and researchers in safety-related domains from our social networks, and
3. ask these people to disseminate information about our survey.

Our sampling procedure can best be described as a mixture of opportunity, volunteer, and cluster-based sampling. The cluster is formed by survey participants from several of these channels. We expect to get a sample stronger than non-probabilistic but, because of a lack of control of the sampling process, weaker than uniformly random.

Sample Representativeness. To check how well our final sample appropriately represents safety practice and its practitioners, the questionnaire measures the classification criteria in Table 1. See [13] for the question used for this.

3.6. Analysis Procedure

This section describes the analysis of the responses, the checking of the working hypotheses, and our tooling.

3.6.1. Analysis of Responses

We use instruments of descriptive statistics [51] such as median (med), mean (µ), variance (var), and frequency histograms to summarise the responses per question.

Half-Open and Open Questions. Some questions of our questionnaire are half-open, that is, we allow to add another answer option by providing an extra scale and a text field, and some questions are open, that is, we only provide a text field.

Particularly, most demographic questions are half-open multiple-choice (MC) questions, that is, they have an extra text field “Other”. We use the answers from this text field to extend and revise the classifications imposed by the given answer options. See Section 4.2 for the results.

Furthermore, we close some of the main questions using qualitative content analysis and coding [52]. For some half-open questions, we extend the statement lists and nominal scales accordingly. The results of this step are shown in Section 4.3 when discussing the questions in the Sections 4.3.1, 4.3.2, 4.3.9, 4.3.11 and 4.3.12.

3.6.2. Hypothesis Analysis and Statistical Tests

We use non-statistical analysis for all base hypotheses for which we directly collect data (Table 5).

For most comparative hypotheses, we apply the MANN-WHITNEY U test [51] (U for short) to check for difference. We use U if the following assumptions hold:

- exactly one LIKERT-type or ordered-categorical dependent variable (DV),
- random division into two groups,
- group members are not paired,
- treatments via independent variables (IV) are already applied,
- group sizes may differ and be small (< 30),
- per-group distributions of the DV may be dissimilar and non-GAUSSIAN.

Let H be a hypothesis and α be the maximum chance of a Type I error, that is, incorrect rejection of the null hypothesis H0. U tries to reject H0 with a confidence of 1 − α. We require p < α for the Type I error p of incorrectly distinguishing two groups of respondents with respect to H. If U succeeds to reject H0 then the support of the desired alternative hypothesis H1 is increased. Failure of U in rejecting H0 (i.e., p ≥ α) denies any conclusion on H from the given data set [53, p. 168]. The medium maturity and criticality of our hypotheses (for an exploratory study) and the medium accuracy of our data (from a survey method) make it reasonable to stick with the typical choice of α = 0.05.

Acceptance Criteria (AC). The criteria in Table 5 describe the aggregation of the question scales in Table 3 to match the hypotheses. These criteria are built from symbols of the kind q(id)(answer option)(scale value(s)) referring to the questions in Table 4. We require med to be non-central to express a large supportive majority. Alternatively, percentage thresholds (e.g. > 25%) express the desired variance or shape of the distribution. In hypothesis tests, we mainly use classification criteria (Table 1) as IVs.

3.6.3. Tooling

We use Unipark6 as a platform for implementing on-line surveys and for data collection (Section 3.5) and temporary storage. For statistical analysis and data visualisation (Section 3.6.2) we use GNU R7 and Unipark. Content analysis and coding takes place in typical spreadsheet applications.

5For example, our construct envisages hypothesis hInsufMeth to refer to Adequacy of means. However, to keep our questionnaire lean, with question qAppOfMeth, we directly measure agreement for one instance of this hypothesis.

6See http://www.unipark.de.

7See https://www.r-project.org.
3.7. Validity Procedure after Survey Planning

In the following, we evaluate the face and content validity of our instrument, and the internal and construct validity of our study. Although, we did not perform an independent pilot study according to [29], we took several measures to assess the validity of our study.

3.7.1. Instrument Evaluation: Face and Content Validity

First, both authors performed several internal walk-throughs to improve the survey design and the data analysis procedure.

Second, along the lines of a focus group, we asked independent persons to complete the questionnaire and to provide feedback via an extra form field in the questionnaire and via email. This dry run took place between 13 and 27 June 2017. We gathered 7 independent responses, from 2 postgraduate research assistants with experience in the survey method, and with experience in safety-critical software, systems, and requirements engineering, 1 master student with industrial work experience in safety-critical systems engineering, 1 IT practitioner and English native speaker, 1 person with a health and safety background, 2 persons with a software engineering background.

The feedback from these respondents resulted in

- an extension and balancing of answer options,
- the alignment of answer scales throughout the whole questionnaire,
- improvement of the nomenclature (terms are now described on the questionnaire page they first appear),
- an extension of open answer fields, and
- linguistic improvements.

These steps helped us to improve questionnaire completeness, consistency, and comprehensibility and reduce researcher bias [13, Sec. 3.3.4].

3.7.2. Internal Validity of the Analysis Procedure

Why would the procedure in Section 3 lead to reasonable and justified results?

U is applicable only if groups are independent with respect to the considered IV. This circumstance can cause problems with MC-questions. For example, with qNotionOfSafety, the same respondents might be in both groups “all data points with choice (c)” and “all data points with choice (a).” Hence, these groups contain data points that can be dependent in a certain but unknown way. For comparisons with U, we reduce this issue by converting the responses to single-choice questions using discriminating features such as the time of the first answer option chosen.

The 7 test data points allowed us to validate our tooling (e.g. R scripts, see Section 3.6.3). Test data points are not included in the final data set.

3.7.3. Construct Validity

Why would the construct (Section 3.1.1) appropriately represent the phenomenon to investigate?

Because of the exploratory nature of our study, the subconstructs and their scales in Table 2 represent the study object as reconstructed from our analyses. The working hypotheses and the questionnaire represent an approximation and a selection of what needs to be measured and tested if we were to investigate this study object (cf. Figure 2) in an explanatory study. For example, we assume that the 10 statements in Figure 18 for question qIntOfSafSec satisfactorily approximate the “interaction of safety and security activities” (i.e., construct Collaboration) and its criticality. Consequently, the scales in Table 2 serve as a reference to the internal validation of our study.

Not being unusual for an exploratory study, several of the hypotheses are relatively weak and, therefore, even if accepted from our collected data, only allow the derivation of restricted conclusions. For example, an accepted hFMsImprSaf (i.e., FMs have a positive impact) reflects very much the personal experience, perception, or opinion of our survey participants. Their view has to be distinguished from the question of actual FM effectiveness. To pursue such a question, we have to refine our research design using the technology acceptance model [54] and controlled field experiments.

Our inquiry of SPs about supportive factors and obstacles in safety decision making does not include safety evidence traceability and management. A future version of our construct and questionnaire should therefore include the criteria examined by Nair et al. [24], De la Vara et al. [26] and Borg et al. [25].

Our construct and instrument are essentially new. However, the Constituents overlap with the construct used in [22] for safety process maturity assessment. Furthermore, Manotas et al. [55] employs a research design analogous to ours, underpinning the appropriateness of our approach to survey engineering practitioners. Despite the drawbacks discussed before, we believe our design is appropriate with respect to the expressive power of the working hypotheses. In summary, this construct can be a helpful guidance in the design of successive studies.

3.7.4. Reliability

A check for test-retest reliability (e.g. changing attitudes of respondents) and alternate form reliability are out of scope of this exploratory study. Hence, we do not plan to ask respondents to answer the questionnaire more than once and we run only one variant of the questionnaire.

4. Survey Results

In this section, we characterise our sample (Section 4.2), summarise the responses (Section 4.3), and analyse our hypotheses (Section 4.4).
Table 6: Safety-related channels we advertised our survey on (sorted alphabetically by category, full list in [13, pp. 92f])

| Channel Type       | Example/References                                                                 |
|--------------------|------------------------------------------------------------------------------------|
| Facebook sites     | E.g. Int. Society of SPs                                                           |
| General panels     | SurveyCircle, www.surveycircle.com                                                 |
| LinkedIn groups    | E.g. on ARP 4754, DO-178, ISO 26262                                               |
| Mailing lists      | E.g. system safety (U Bielefeld, formerly U York)                                 |
| Newsletters        | GI requirements engineering                                                         |
| Personal websites  | E.g. profiles on Twitter, LinkedIn, Xing                                            |
| ResearchGate       | Q&A forums on www.researchgate.net                                                  |
| Xing groups        | E.g. safety engineering                                                             |
| Other channels     | E.g. board of certified safety professionals                                        |

4.1. Survey Execution: Sample Size and Response Rate

For the collection of data from the survey participants, we

1. advertised our survey over the channels in Table 6 and
2. personally invited > 20 persons.

The sampling period lasted from 1 July 2017 til 25 September 2017. In this period, we repeated step 1 up to three times to increase the number of participants. The Unipark tracking data shows that LinkedIn groups, ResearchGate, Twitter, and mailing lists were effective in soliciting respondents, however, it is incomplete and, hence, does not disclose which channels were most effective.

After 565 views of the questionnaire, our final sample contains \( N' = 124 \) (partial) responses with \( N = 93 \) completed questionnaires and \( N = 91 \) (73%) complete\(^8\) data points. Figure 3 depicts the distribution of responses over time. According to our questionnaire tool, respondents spent 20 minutes on average to provide complete data points, 50% spent within 14 and 24 minutes time.

Given the numbers of members for some channels we used (e.g. for LinkedIn groups), we estimate the return rates of responses per channel to range from 0.1 to 5%.

From the sub-groups, we build from the sample according to our classification criteria, the smallest we like to reason about below are of the size of around 15.

Appendix A provides a detailed enumeration of data summaries (i.e., number of answers per option) for all questions.

4.2. Description of the Sample

We describe our sample in the following and estimate the extent to which it represents (Section 3.5) the population of SPs. For each classification criterion according to Table 1, we provide a chart or we name the up to 10 most frequently occurring answers, ordered by frequency. Percentages (%) right of the bars indicate the fraction of the 93 completed questionnaires, shown in parentheses the (N)umber of respondents who chose the corresponding option. Note that most of the classification questions allow MC answers (cf. Table 1 and [13, pp. 56ff]).

Educational Background. Figure 4 summarises the educational background of all respondents:

- **Computer scientists** include software engineers and computer engineers
- **Electrical and electronics engineers**
- **Safety scientists** include safety engineers, occupational safety practitioners, health and safety practitioners, human factors engineers, ergonomics engineers
- **Mechanical and aerospace engineers**
- **Systems engineers** include poly-technical systems engineers, information systems engineers, business technologists, engineering business administrators, engineering project managers
- **Physicists and mathematicians**
- **Other discipline** includes chemists, biochemists, civil engineers, language scientists

Application Domains. Figure 5 summarises the application domain of all respondents where “aerospace” includes space telescopes; “industrial processes and plant automation” includes manufacturing, chemical processes, oil and gas, energy infrastructure, and small power plants; “railway systems” includes railway signalling; “construction and building automation” includes civil engineering applications; and “other domains” includes food safety, biological safety, research and development, and environment, health, and safety preparations.
Level of Experience. Figure 6 indicates that our sample of SPs is moderately balanced across all experience levels.

Familiarity with Standards. Figure 7 provides an overview of safety-related standards our respondents are familiar with (distinguished by generality or by application domain): Standards from aerospace (e.g., ARP 4761, DO-178, DO-254), generic standards (e.g., ISO 61508, DIN VDE 0801) automotive (e.g. ISO 26262), machinery (e.g. ISO 13849, 25199, DIN EN 62061, MRL 2006/42/EG), military (e.g. MIL-STD 882, UK Def Std 00-55), railway (e.g., CENELEC EN 50126, 50128, 51029, 62061), power plants (e.g. IEC 60880, 61513, 62138, 60987, 62340, IEC 8000), and medical devices (e.g. IEC 80001, ISO 14971, AAMI/UL 2800). 14 participants were neither familiar with any of the given standards nor did they specify other standards.

Familiarity with Methods. Figure 8 shows the familiarity of our respondents with prevalent concepts of safety analysis and the corresponding classes of methods, techniques, or notations:10 For example, FMEA, FMECA, or FMEDA to assess failure mode effects; HazOp studies, ergonomic work analysis and intervention methodology to assess hazard operability; STAMP-based methods for hazard (STPA) and accident (CAST) analysis; FTA, FMEA, PHA, or FTA to assess risk at a functional or abstract design level; common cause (CCA) or common mode (CMA) analysis to include dependencies and interactions; fault injection and property checking as techniques of automated validation and verification (V&V); STRIDE or CORAS to assess and handle security threats; bidirectional methods such as Bowties or cause-consequence analysis; MARKOV chains for probabilistic risk analysis, and GSN and SACM to build assurance cases. For “Other”, our participants mentioned a variety of approaches (no more than twice): 5S, 5W, CASS, coexistence analysis, FRAM, HazRAC, HEART, HRA, MTA, (O)SHA, SAR, SCRA, SHARD, SSA, Poka Yoke, prognostic analysis, WBA, ZHA, ZSA.

Only 4 respondents state familiarity with methods to assess and handle security threats. 15 respondents neither checked any of the given methods nor did they specify other methods that are relevant in their safety activities.

Geographical Regions. DE (24.3%), UK (16.4%), US (15.3%), AU (6.2%), FR (5.1%), IT (3.4%), CA (3.4%), CN (2.8%), and CH (2.8%).

Native Languages. Figure 9 provides an overview of the languages spoken by the respondents.

Working Languages. Figure 10 provides an overview of the languages used at work by the respondents.

Safety-related Roles. In Figure 11, the term practitioner includes the profile of an engineer and a manager. Regarding engineering disciplines and domains, “safety practitioner” includes engineers or managers in system safety, functional safety, or in other safety domains as well as technology risk managers in general; “software practitioner” includes developers, architects, and tool developers; “systems practitioners” includes system analysts.

---

10Abbreviations are described in Table B.9 in Appendix B.
and system architects; “health & safety practitioner” includes occupational safety practitioners, human factors engineers, and ergonomists; and “V & V practitioner” includes test and assurance practitioners. For “Other”, our respondents include a civil engineer, a project manager, a method engineer, and a maintainability engineer.

Regarding responsibility profiles, the category “Consultant / Assessor” includes independent evaluators, auditors, regulators, and inspectors dealing with safety certification.

### 4.3. Summary of Responses

In this section, we summarise the responses to the questions in Table 4.

**Guide to the Figures.** The following text and figures complement each other. For Likert-type scales, we use centred diverging stacked bar charts as recommended by [56]. med denotes the median and “ex” indicates the number of excluded data points per answer option.

#### 4.3.1. qValueOfKnow: Value of Knowledge Sources

Figure 12 shows that, among the knowledge sources we asked our participants to rate, expert opinion, previous experience in safety-related projects, and case reports represent the three highest valued knowledge sources used in safety activities and safety decision making. Management recommendations turn out to be the lowest valued knowledge source.

The following knowledge sources, or resources in more general, were additionally mentioned to be of very high or high value:

- Four respondents referred to the concept of adversarial thinking, mentioning “creative mind”, “imagination”, “analysis capability,” and “acceptance of human fallibility.” Three respondents pointed to the concept of domain expertise and experience, mentioning “gut feel”, “subject matter knowledge of the application,” and “real work and related problems in reference situations...” Furthermore, they mentioned education, specification documents

Figure 9: Native Languages and concepts (frequency, MC)

![Figure 9](image)

Figure 10: Working Languages and concepts (frequency, MC)

![Figure 10](image)

Figure 11: Safety-related Roles (frequency, MC) split into disciplines (top) and responsibility profiles (bottom)

![Figure 11](image)

Figure 12: qValueOfKnow (N = 97): Value of knowledge sources – Of how much value are specific knowledge sources for safety decision making? and tools (e.g. “use of SPARK”), independent assessment, in-service monitoring logs, and previously certified similar systems.

#### 4.3.2. qImpOfConstr: Constraints on Safety Activities

According to Figure 13, inexperienced safety engineers (g) and erroneous hazard analyses (e) gained the most ratings in the category “significant negative impact on safety activities.” Postponed safety decisions (c) achieved the largest consensus. Vague safety standards (f) constitutes the bottom of this ranking but is still rated with medium or high negative impact by the majority of respondents.

The following factors (i.e., process constraints and issues) were additionally mentioned to have high negative impact on safety activities:

- Eight respondents broach the issue of missing management expertise and support: “Lack of education of managers in need for safety” identifies one respondent from the oil and gas industry. Another one states that there is a “general perception that safety is only paper work”
and perceives a “lack of safety knowledge within management.” One practitioner was even pointing to a “lack of general safety culture.”

Three participants criticise that the degree of collaboration is too low: They perceive a “lack of system level engineering experience” as well as “soloed working practices without a clear view of an integrated safety concept” and that the organisation is “minimising the involvement of safety process/engineers into the development process.”

Regarding incomplete or inadequate hazard lists, respondents mention “unrecognisable hazard domains” and “imagined safety cases not based on real workers experience.” Along with that, one practitioner mentions the issue of “poorly defined requirements”: Such requirements, when coming from upstream, are known to have a negative effect on many downstream engineering activities. Conversely, inadequate hazard lists resulting from such activities can again have a negative impact on downstream subsystem requirements specification.

Regarding compliance with norms, one respondent was criticising the “transfer of concern from assessment to compliance,” in other words, compliance bias. Two others are broaching the opposite phenomenon of compliance ignorance, mentioning “general ISO 26262 standard ignorance” and a “lack of understanding of regulatory framework.”

Furthermore, according to another participant’s experience there is “too much faith in testing” and “reluctance to use formal methods.”

### 4.3.3. **qImpOfEco**: Influence of Economic Factors

More than a third (36%) of the survey participants share the view that economic factors often strongly influence the way how hazards are handled, about half of them (48%) think that such influence happens rarely or occasionally (median), and for 9% such influences are not recognisable.

Figure 13: **qImpOfConstr** (N = 100): Negative impact on safety activities – To which extent do specific process constraints and issues negatively impact safety activities?

![Figure 13: qImpOfConstr](image)

Figure 14: **qAdeqOfMthStd** (N = 102): Adequacy of methods and standards – Regarding a specific application domain, how adequate are applicable safety standards and methods in ensuring safety?

![Figure 14: qAdeqOfMthStd](image)

Figure 15: **qAppOfMeth** (N = 97): Applicability of methods – The application of conventional techniques (e.g. FMEA and FTA) has become too difficult for complex applications of recent technologies.

![Figure 15: qAppOfMeth](image)

4.3.4. **qAdeqOfMthStd**: Adequacy of Methods and Standards

According to Figure 14, traffic control (f) and medical and healthcare applications (e) are most often believed to be supported by adequate methods and standards. However, for all domains, at least 50% of the respondents think that the available means are only slightly or not at all adequate for safety assurance. This question exhibits a relatively large number of donk-answers.

4.3.5. **qAppOfMeth**: Applicability of Methods

The nand-median in Figure 15 shows that there is no tendency or no clear consensus among respondents on whether or not conventional methods have become too difficult to apply in current applications.

4.3.6. **qPosImpOfFMs**: Positive Impact of Formal Methods

The median of “medium impact” in Figure 16 indicates a consensus among the participants on that the use of FMs might have a positive impact on the effectiveness of safety activities. However, we only have a low number of responses resulting from missing answers and we excluded donk-answers.

4.3.7. **qImpOfSkills**: Improvement of Skills

According to Figure 17, SPs agree moderately (39%) to strongly (54%) with the requirement to adapt their professional skills to new technologies. However, significantly
less consensus was achieved among the respondents on whether junior SPs should learn from accident reports.

4.3.8. qIntOfSafSec: Interaction of Safety and Security

The high moderate and strong agreement in Figure 18 indicates that most of our participants perceive interactions between safety and security as critical.

SPs clearly agree on that interaction between safety and security practitioners during requirements engineering and system assurance rarely occurs (g,f). Furthermore, clear agreement is achieved for the “negative influence of a lack of collaboration (between safety and security engineers)” (h,i) and for the “positive influence of such a collaboration” (j) on safety activities. However, we acknowledge 7% of disagreement with the “requirement of ultimate IT security for safety.”

No clear consensus is achieved regarding the dependence of security on safety (b,d). As opposed to that, respondents agree on the dependence of safety on security (a,c,e).

4.3.9. qNotionOfSafety: Notion of Safety

The multiple-choice answers in Figure 19 show that many participants seem to be reluctant to associating cost/benefit schemes with management decision making in system safety (a,b). Accordingly, many responses indicate that safety is treated as a cost-independent necessity (c). However, 51 (32%) responses were given to the view of safety as an “important, yet secondary, and tedious mandated issue” (d,e).

Beyond the five given answer options, the notions of safety additionally given by our respondents range from a “huge effort generating source”, a “marketing gadget”, a...

4.3.10. qPrioOfSafety: Priority of Safety

From Figure 20, we can see a clear consensus of the respondents for all given options (a–d). Particularly, increased priority of safety decisions (a) and defined safety processes (d) positively contribute to the efficiency of safety activities. The Sections 4.3.11 and 4.3.12 provide more details on the factors believed to increase the efficiency and effectiveness of safety activities as well as dual factors assumed to decrease the efficiency and effectiveness thereof. Along the way, comparably many SPs (17%) do not offer any agreement on authority (c).
4.3.11. qEffRoleOnJob: Effect of Role Model on a SP's Job

We asked our respondents to comment on whether and how their job is affected by a clear definition of their role, if any, in their organisations and application domains.

Apart from 5 *dnk*-answers, we received 56 answers saying “yes” and, thus, stating that the role of a SP is clearly defined. These SPs perceive or expect the following positive consequences on their job (frequency given in parentheses, in descending order): Clear role definitions . . .

- have a general positive impact on a SP's activities (24),
- lead to clear responsibilities, authority, and escalation routes (13),
- allow good integration of safety activities into the surrounding system life cycle processes (6),
- can make the achievement of compliance easier (1), and
- let SPs maintain autonomy or independence to carry through their most critical activities (1).

However, our study participants report on the following negative effects on their job: Clear role definitions . . .

- make engineers entirely push away safety-related responsibilities as a consequence of separating teams into safety and non-safety co-workers (2),
- lead to complex process definitions (1),
- get rather independent SPs exposed to company-wide resource and risk management (1), and
- impose a wrong focus or unnecessarily constrain a SP's tasks (1).

Moreover, 30 participants responded with a “no” and, hence, state that the role of a SP is not clearly defined. These SPs consider or expect the following positive consequence on their job: Unclear role definitions . . .

- may promote more freedom to act, for example, to develop and employ new and more effective safety approaches (3).

However, our respondents also perceive several negative effects on their job: Unclear role definitions . . .

- can entail unclear or wrong responsibilities as well as limited authority, autonomy, and space for discretionary activity (9),
- promote unclear, one-sided, or late decision making, in the worst case, rushed processing of checklists (6),
- have a general negative impact on a SP’s tasks (4),
- can lead to disintegrated conceptions of safety, separated communities with a lack of communication and coordination, promoting unnecessarily confined decisions (3),
- can decrease the appreciation of a SP’s analysis capabilities (2), and
- increase the risk of unqualified personnel assuming the role of a SP (2).

4.3.12. qEffNotionOnJob: Effect of Safety Notion on a SP's Job

We asked our respondents to comment on whether and how their job is affected by a predominant notion of safety (qNotionOfSafety), if any, in their organisations and application domains.

Apart from 9 *dnk*-answers, we received 74 answers indicating a “yes” and, hence, stating that the notion of safety has an effect on their job: 10 respondents do not provide a specific comment. The others argue from several notions of safety they have perceived in their environments. Below, we provide answer frequencies and cite a few answers underpinning the summary statements.

Non-supportive Notions of Safety. 24 participants describe their experiences with a non-supportive, misunderstood, or underrated safety culture. They report that . . .

...SPs have difficulties to argue their findings (9): “Right now there is no ability to have the safety requirement override standard functional requirements.” – “1. Our job always gets delayed and we are the last to get the inputs. 2. Non-safety engineers always try to justify or avoid the suggestions/findings. 3. It is difficult to sell safety culture to non-safety engineers/managers.” – “I have to spend extra time explaining that safety is not about compliance or implementing controls.” – “As for now safety has not the degree of importance to support testing views and arguments against system designers and management.”

...SPs suffer from late decision making (5): “If I am not allowed to do my job early in the process (requirements stage), safety becomes more costly and I as a safety practitioner am viewed as a late check in the box to get through a program rather than an integral part of a design team.”
...SPs’ activities have no lasting value (1): “The safety practitioner is neither equipped, nor capable of making the decisions needed for a higher level of safety. Being mostly policemen, enforcers and rule designers, little if any of their contributions have any meaningful or lasting value.”

Supportive Notions of Safety. 20 respondents describe their experiences with or their view of a supportive or highly-valued safety culture. They report that...

...SPs’ findings are important and heard (6): “My job is important because safety is valued and considered necessary.” – “Most people in my organisation understand the importance of safety. This is positive.” – “There are not many people who practice safety, since it is a tedious job. So we are highly valued.”

...SPs are properly included in the process (1): “Safety is fundamental to the work we do and is ingrained into our processes in such a way that its impossible to ignore. While it makes jobs harder with much more analysis and review processes and every stage of the product’s development, we know its vital.”

Other Notions of Safety. 9 SPs describe an ambivalent picture, saying that it depends on individual projects whether their jobs are negatively or positively affected: “Safety at the last two places I worked is a check box activity at best. Other places I’ve worked it was started early in the pre-design phase. Starting early is more cost and schedule effective with a better end product.”

5 SPs refer to a regulation-driven notion of safety: “Positively affected. In aerospace, safety is part of fundamental engineering principles, so the process is embedded in systems engineering and does not get left out.”

From a budget- or schedule-driven perspective, respondents (4) observe that “the budget for tools and training is never enough” and that “resources, budget, support depend on the view/culture of safety.”

Finally, 12 respondents claim, by saying “no”, that the notion of safety does not have any effect on their jobs.

4.3.13. qUndesiredEv: Role of Undesired Events for Safety

Figure 21 shows a clearly disagreeing response on whether lack of failures reduces the need for carrying through safety activities (a). We have a more ambiguous agreement on whether safety implies reliability (e), that is, on whether having assured the safety of a system usually includes having also assured the reliability of a system. Moreover, known and reported accidents seem to be important for the argumentation of the need for safety (b,c). However, the agreement on whether a “lack of accidents weakens arguments for the need of safety” (d) varies more.

4.3.14. qValOfContrib: Value of SPs’ Contributions

According to Figure 22a, the majority of respondents perceives their role in the system life cycle as highly valu-
Figure 23: qImpOfExp (N = 96): Role of experience – Specify your level of agreement with 3 statements about the role of experience and adversarial thinking in safety activities.

Finding 1. The responses suggest that safety mainly depends on expert opinion and project memory.

Finding 2. Standards in the considered high automation domains seem to be inadequate.

Finding 3. From our data, we are not able to provide a clear general picture about the adequacy of means.

Finding 4. Among informed respondents, formal methods are believed to be beneficial.

5. Discussion

We interpret the responses (Section 5.1), draw a relationship to existing evidence (Section 5.2), and critically assess the validity of our study (Section 5.4). From these discussions, we derive our conclusions in Section 6.

5.1. Interpretation of the Results and Findings

The following discussion takes into account the hypothesis analysis and test results summarised in Table 7. Details about hypotheses and questions referred to in the text can be derived from Tables 4 and 5.

5.1.1. Findings for RQ1: Means of Work in Safety Practice

Hypothesis hExpDecides. The support of hExpDecides should not be surprising as it mirrors a rather typical situation in many engineering disciplines and projects. However, relying too much on knowledge of experts can in the worst case go along with relying on a single point of failure of an organisation. Moreover, relying too much on experience from similar projects can unfortunately go along with wrongly transferring former conclusions (i.e., project memory) and not updating them correspondingly.

Hypothesis hInsufMeth. Because of overlapping Means, the rejection of hInsufMeth stands in conflict with the support of hInsufStd. On the one hand, we see a slight tendency towards the first author’s experience from interviews [5, 57, 4] suggesting hInsufMeth to be a justified hypothesis. On the other hand, ambiguous agreement was given to “have become too difficult” or, more generally, to “have become inadequate.” Asking for agreement in question qAppOfMeth should have been substituted by asking for the level of Adequacy. However, we believe it is safe to interpret the respondents’ agreement that “available standards and methods have become too difficult” as “they are challenging to apply.” After all, we conclude that this construct should better be measured by several questions to get more informative and reliable results.

Hypothesis hFMsImprSaf is supported. The low number of valid responses to question qPosImpOfFMs certainly weakens the interpretation of the support of hFMsImprSaf. Both, the question qPosImpOfFMs as well as the notion of a formal method are very abstract. Moreover, the classification questions provide only little knowledge about our respondents’ experience with FMs.

Finding 1. The responses suggest that safety mainly depends on expert opinion and project memory.

Finding 2. Standards in the considered high automation domains seem to be inadequate.

Finding 3. From our data, we are not able to provide a clear general picture about the adequacy of means.

5.1.2. Findings for RQ2: Impact of Process Factors

Hypothesis hLoResLoSaf is rejected. Few respondents to question qEffNotionOnJob experience a lack of resources for safety activities. This is consistent with the data checked for the AC of hLoResLoSaf. Although the responses suggest that the implication lack of resources has negative impact on safety might hold, the antecedent of this hypothesis is not broadly supported.

Certainly, this finding requires another study with a more specific research design.
Table 7: Results of hypotheses analysis for hExpDecides to hPosSelfImg and hypotheses tests for hExp:DivGTSing to hInsufMeth:AutoGTAero. Legend: AC... acceptance criterion

| Hypothesis | From the responses to... | ...we conclude that... |
|------------|-------------------------|-----------------------|
| hExpDecides: Dependence on expert opinion | qValueOfKnow (Section 4.3.1) and qImpOfExp (Section 4.3.16) | our AC is fulfilled. |
| hLoResLoSaf: Resources govern performance of SPs | qImpOfConstr (Section 4.3.2) and qImpOfEco (Section 4.3.3) | the qImpOfEco-part of our AC is not fulfilled. |
| hInsufStds: Inadequate means in high-automation | qAdeqOfMthStd (Section 4.3.4) | our AC is fulfilled. |
| hInsufMeth: Low method adequacy | qImpOfEco (Section 4.3.3) | the qImpOfEco-part of our AC is not fulfilled by the nand-median. |
| hFMsImpIrSa: Positive impact of formal methods | qAppOfMeth (Section 4.3.5) and qPosImpOfFMs (Section 4.3.6) | our AC is fulfilled. |
| hSPsAdptSkl: Necessity of skill adaptation | qImpOfSkills (Section 4.3.7) | our AC is fulfilled. |
| hSafTsCost: Safety is a cost-benefit question | qNotionOfSafety (Section 4.3.8) | none of the qDesiredEv-parts of our AC are fulfilled. |
| hLoCollLoSaf: Benefit of safety-security interaction | qNotionOfSafety (Section 4.3.9) | the qNotionOfSafety-a-part of our AC is not fulfilled. |
| hHiPrioHiSaf: Benefit of safety-as-a-priority | qImpOfEco (Section 4.3.3) | our AC is fulfilled. |
| hSafTsRel: Safety is a special case of reliability | qImpOfEco (Section 4.3.3) | none of the qDesiredEv-parts of our AC are fulfilled. |
| hSafTsValued: High contribution to life cycle | qCoWorkers (Section 4.3.15) and qValOfContrib (Section 4.3.14) | the qCoWorkers-part of our AC is not fulfilled. |
| hPosSelfImg: High contribution (self-image) | qCoWorkers (Section 4.3.15) | our AC is fulfilled. |

From the comparison of... ...we conclude that...

| Hypothesis | From the responses to... | ...we conclude that... |
|------------|-------------------------|-----------------------|
| hExp:DivGTSing: Benefit of diverse expertise | senior SPs with junior SPs (from responses to qImpOfExp, Section 4.3.16) | our AC is fulfilled. |
| hValue:SenLTJun: Assoc. of expertise & value | senior SPs with junior SPs | with \( p = 0.15 \), our AC is not fulfilled. |
| hAdapt:SenGTJun: Assoc. of expertise & skill adaptation | senior SPs with junior SPs | with \( p = 0.092 \), our AC is almost fulfilled. |
| hAdapt:AutoGTAero: Assoc. of standards & skill adaptation | SPs using automotive standards with SPs using aerospace standards | with \( p = 0.09 \), our AC is almost fulfilled. |
| hInsufMeth:EngDifSci: Assoc. of profession & inadequate means | engineering-focused SPs with research-focused SPs | with \( p = 0.12 \), our AC is not fulfilled. |
| hInsufMeth:AutoGTAero: Assoc. of standards & inadequate means | aerospace-focused SPs with aerospace standards | with \( p = 0.24 \), our AC is not fulfilled. |

**Finding 5.** Our data suggest that resources occasionally but not typically govern SPs' performance.

However, by weakening hLoResLoSaf, we can acknowledge the “often” third of SPs showing a situation demanding for reaction in the community.

**Hypothesis hSafTsCost is rejected.** We identify a weak positive association: Safety is most frequently viewed as a cost-independent necessity (qNotionOfSafety-c, hSafTsCost) and the median of qImpOfEco (hLoResLoSaf) lies at economic factors rarely or occasionally influence safety. So, for hSafTsCost, the many positive responses to the options (c,d,e) underpin the view of safety as a cost-independent factor in management decision making. We consider this to be positive but like to stress the need of an in-depth explanatory study to confirm or refute this finding.

**Finding 6.** Our responses suggest that safety is not typically a question of cost-benefit.

Hypotheses hLoCollLoSaf is supported. First of all, our data supports hSafBySec which states that safety assurance strongly depends on security assurance. Interestingly, for hLoCollLoSaf, SPs agree on both that...

**Finding 7.**...a lack of collaboration or interaction downgrades the performance of safety activities (qIntOfSafSec-h,i), and

**Finding 8.**...interaction between safety and security practitioners rarely occurs in requirements and assurance activities (qIntOfSafSec-f,g).

We consider this issue worthwhile to be monitored. Apart from desirable interactions at an organisational level, potential dependence of security on safety (qIntOfSafSec-b,d) is less obvious to our respondents than potential dependence of safety on security (qIntOfSafSec-a,c,e). While the latter is comparatively well known, the former is more difficult to grasp. Our data shows this ambiguity but does not explain it.
Finding 9. Overall, collaboration of safety and security experts is clearly viewed as beneficial.

Hypothesis $h_{Exp:DivGTSing}$ is supported. Although the three propositions in Figure 23 seem obvious, we included them in our questionnaire to confirm that such occasionally important assumptions are actually made by SPs ($h_{Exp:DivGTSing}$). For these assumptions to be formulated as a hypothesis and tested accordingly, a further investigation would be necessary. Hence, the support of $h_{Exp:DivGTSing}$ is not very informative on its own but backs the support of $h_{ExpDecides}$.

Finding 10. Diverse expertise is perceived as beneficial for SPs.

Hypothesis $h_{Adapt:SenGTJun}$ vs. $h_{Adapt:AutoGTAero}$ and $h_{SPsAdaptSkls}$. Among the comparative hypotheses, only $h_{Adapt:SenGTJun}$ and $h_{Adapt:AutoGTAero}$ are close to being supported with $p = 0.052$ and $p = 0.09$. The result for $h_{SPsAdaptSkls}$ is unsurprising because senior experts are professionals with longer experience and might have witnessed training activities in their field more often than junior SPs. However, the small difference between both groups gives rise to the conjecture that senior experts would avoid outdated skills as much as junior professionals would. An almost supported $h_{Adapt:AutoGTAero}$ gives rise to the conjecture that in automotive, currently in demand of improvement of their safety practices, SPs spend corresponding effort on skill improvement.

Finding 11. The improvement of skills towards new technologies is generally agreed among respondents.

5.1.3. Findings for RQ3: Perception of Safety Practice

Hypothesis $h_{SafIsRel}$ is rejected. Similarly, we perceive the results for $h_{SafIsRel}$ as positive because the issue of “confusing safety with reliability” raised in [49, p. 7, Assumption 1] can at least not be confirmed from the analysis of our responses. In fact, we observe an opposite tendency from our sample and assume this to be the effect of those SPs having been trained on that issue.

Finding 12. It is generally justified to not believe in the hypothesis “safety is equivalent to reliability.”

From the responses to qUndesiredEv-a, we derive that assured reliability of a system does not reduce the need for safety activities. Consequently, these responses do not give rise to believe in the hypothesis reliability implies safety. However, we might sometimes expect to see agreement on the hypothesis safety implies reliability ($q_{UndesiredEv-e}$). Likewise, our responses are ambiguous in that case. The most reasonable explanation for this ambiguity is that we missed to clearly explain what such implications exactly mean when used as answer options. Moreover, $h_{SafIsRel}$ is not backed by redundant data. The data gathered from qUndesiredEv makes it hard to draw a strong conclusion. To back a “true extension” of reliability—that is, safety carries features essentially different from reliability or, even more, safety is independent of reliability—we should have asked questions like “Does reliability imply safety?” with an expected median of “disagree”.

In conclusion, our data gives rise to the reasonable belief that safety and safety activities are less dependent on issues of system failures than on the more general issues of system accidents.

Finding 13. From our responses, we cannot further characterise the relationship of safety and reliability.

Questions qEffRoleOnJob and qEffNotionOnJob. 56 respondents state that their role is clearly defined. 37 perceive positive impacts on their activities, particularly, fostering clear responsibilities, authority, and escalation routes. 30 respondents state that their role is not clearly defined. 15 of them perceive negative impacts in form of unclear responsibilities, limited authority, autonomy, and space for discretionary activity as well as unclear or late decision making (Section 4.3.11).

Finding 14. The role of a SP is often not clearly defined and SPs experience negative impacts from this.

24 participants experience a non-supportive, misunderstood, or underrated safety culture. As opposed to that, 20 respondents perceive a supportive or highly-valued safety culture. 9 persons provided an ambivalent picture of safety culture, stating that they have gathered contrasting experiences (Section 4.3.12).

Finding 15. SPs perceive to a similar extent both, supportive and non-supportive notions of safety.

Hypothesis $h_{PosSelfImg}$ is supported. While responses to qValOfContrib support $h_{PosSelfImg}$, the frequent indication of “medium value”, particularly for qCoWorkers, suggests that some SPs might either not be convinced of the role, their profession, or even unsatisfied with their tasks and their job profile. Section 4.3.11 provides some explanation for such a dissatisfaction coming from an unclear role definition and Section 4.3.12 delivers an explanation from a non-supportive safety culture. However, for a solid conclusion, this indication has to be investigated in more detail by further studies.

The perception of an SP’s role and contribution by non-safety co-workers slightly differs from how SPs perceive their own role. This might not be too surprising because qValOfContrib and qCoWorkers redundantly measure fragments of a participant’s self-perception.

Finding 16. SPs seem to be self-confident about their contribution.
Table 8: Overview of main findings from hypothesis analysis

| RQ1: Which means do SPs typically rely on? How helpful are those means to them? | RQ2: Which typical process factors have influence on SPs’ decisions & performance? | RQ3: How do SPs perceive and understand their role in the process or life cycle? |
|---|---|---|
| F1: The responses suggest that safety mainly depends on expert opinion and project memory. | F5: Our data suggest that resources occasionally but not typically govern SPs’ performance. | F12: It is generally justified to not believe in the hypothesis “safety is equivalent to reliability.” |
| F2: Standards in the considered high automation domains seem to be inadequate. | F6: Our responses suggest that safety is not typically a question of cost-benefit. | F13: Our responses do not offer specific insights on the relationship between safety and reliability. |
| F3: From our data, we are not able to provide a clear general picture about the adequacy of means. | F7: A lack of collaboration or interaction downgrades the performance of safety activities. | F16: SPs seem to be self-confident about their contribution. |
| F4: Among informed respondents, formal methods are believed to be beneficial. | F8: Interaction between safety and security practitioners rarely occurs in requirements and assurance activities. | F14: Their role is often not clearly defined and SPs experience negative impacts from this. |
| | F9: Collaboration of safety and security experts is clearly viewed as beneficial. | F15: SPs perceive to a similar extent both, supportive and non-supportive notions of safety. |
| | F10: Diverse expertise is perceived as beneficial for SPs. | |
| | F11: The improvement of skills towards new technologies is generally agreed among respondents. | |

5.2. Relation to Existing Evidence

In Table 8, we summarise our findings and, below, we compare them with findings from related studies.

Graaf et al. [18] identify legacy incompatibility, lack of maturity, and additional complexity of new methods, languages, and tools as three major obstacles to the early or timely adoption of such means. Similar obstacles were observed in software testing by Kasurinen et al. [19]. These observations are consistent with F1 that knowledge from previous projects has the strongest influence.

Martins and Gorschek [12] observe a lack of evidence for the usefulness and usability of new approaches from safety research. Their observation is not in conflict with finding F4, because SPs can perceive usefulness of new FMs independent of evidence. The authors perceive a dominance of conventional approaches in practice which is again consistent with finding F1. Furthermore, they observe a lack of studies that investigate how to improve the communication process throughout the life cycle. F9 indicates that such studies would be of interest to practitioners.

Chen et al. [21] observe that assurance cases can improve cross-disciplinary collaboration but are missing tool support and experienced personnel. We believe that a lack of research transfer and training could explain the contrast to finding F11, given that assurance cases are seen as a new method by SPs.

Borg et al. [25] and De la Vara et al. [26] clarify finding F3 at least for the specific case of change impact analysis in safety practice.

McDermid and Rae [10] could find no satisfactory explanation to their observation that systems got “so safe despite inadequate and inadequately applied techniques.” However, their assumption is orthogonal to finding F2, contrasts finding F3 and certainly emphasises the need for further empirical research. The lack of consensus on how to combine case-based [20] and compliance-based [22] assurance underpins this lack of clarity on the adequacy of means.

Finding F1 supports the observation of Nair et al. [24] that expert judgements and checklists are among the most frequently used references to assess safety arguments and evidence (see Figure 12). F1 is also shared by Rae and Alexander [34] who conclude that critical aspects of safety analysis (e.g. identifying hazards, estimating risk probability and consequence severity) often rely on expert opinion. Moreover, F1 underpins two out of Wang’s and Wagner’s [23] top ten identified decision making pitfalls.

Leveson [49] observes that safety is pervasively confused (or assumed to correlate) with reliability. The data for the findings F12 and F13 support her conclusion in parts, but the consensus of our responses suggests that there is broad awareness that safety and reliability are first of all two distinct properties of a system.

In summary, we found related supportive and contrasting evidence regarding most findings for RQ1, RQ2, and RQ3.

5.3. General Feedback on the Survey

The last page of our questionnaire contains a text field to leave general comments, for example, an overall opinion, on our survey.

One issue, our survey participants criticised, pertains to the scope and the terminology used in the questionnaire:

The respondents noted that the inquiry is general and does not account for the diversity of safety practices in various industries. Some questions rely on a particular interpretation of safety practice leaving assumptions implicit and risking to get in conflict with other views of system safety, for example, “safety by introduction of controls” versus “safety assurance and assessment.” Moreover, some of the questions are hard to answer because of a lack
of standardised terminology across domains and because of missing topics, for example, legal safety requirements and regulations, human operators, socio-technical systems were not mentioned.

Although this is justified critique, we found it hard to arrive at a terminology and at a level of detail suitable for all SPs while keeping our construct lean (Section 3.1). After several iterations and an email-based focus group, we finalised the questionnaire to be released.

When designing our questionnaire, we were driven by specific not necessarily related findings from previous studies. Moreover, we had to prioritise and cut the question catalogue to stay within a maximum duration of about 30 minutes, an amount of time we believe to be affordable by the participants.

Except for qAdeqOfMthStd and qPosImpOfFMs, the acceptably low number (<10%) of dk? responses indicates that most respondents did not seem to struggle with answering most of the questions. However, frequent nand? responses indicate difficulties in deciding on the given answer options (see, e.g. qAppOfMeth).

Another issue raised by our respondents deals with the survey method and design we applied:

Some questions include bias, drive one to answer in a particular way and solicit a specific support. Likert-scales impose an abstraction with the risk to deny more accurate answers such as “I often highly agree and sometimes I strongly disagree.” Moreover, Likert-scales should be substituted by open questions more appropriate for exploratory studies where the construct is not known or (entirely) fixed beforehand.

On the one hand, we have gained good knowledge about the construct from previous studies and, on the other, we provided several possibilities to give open answers and, in fact, present results from their qualitative analysis (e.g. in the Sections 4.3.1, 4.3.2, 4.3.9, 4.3.11 and 4.3.12). More open questions reduce the risks of bias and constrained data acquisition. However, it is worth noting that, as opposed to interviews, too many open answers in large-scale questionnaires can also be demanding for the respondents and, thus, lead to a high number of partial data points.

5.4. Validity Procedure after Survey Execution

Here, we assess our survey design with respect to internal and external validity as well as reliability [53, 58].

5.4.1. Internal Validity

To reduce internal threats to validity, we performed an a-posteriori cross-validation with recommendations on questionnaire-based surveys in the software and systems engineering domain [59]. Section 5.3 discusses further arguments for internal validity as a response to the general feedback on our survey. Additionally, the everyday use of English among the majority of survey participants (Native Languages) supports the accuracy of a large fraction of the data points.

5.4.2. External Validity

To which extent would the procedure in Section 3 lead to similar results with different samples?

Our sampling procedure is network-guided and, hence, not uniformly random [51]. However, on the one hand, from Section 4.2, our sample varies over the scales of all classification criteria (Table 1). This variation limits potential deficiencies of our sample resulting from an overlap of the summer holiday season with our sampling period. On the other hand, regarding the notion of safety culture, our sample might be biased towards the more frequently occurring backgrounds, domains, and geographical regions (Section 4.2). However, the lack of evidence for hInsufMeth:EngDifSci (i.e., practitioners differ from academics in their view of inadequacy of means) reduces the extent to which the participation of researchers biases the results towards a one-sided academic viewpoint.

According to Figure 11, 19 out of 124 respondents stated that they have been working on safety-related topics as a researcher in academia, that is, the role or responsibility profile which we associate the least of all with genuine practical experience. Only 4 of them declared to be solely academic researchers. 8 stated to be SPs, too; 7 have also done research in industry; 11 have worked as software, systems, requirements, reliability, or health & safety practitioners in addition. This again strengthens our belief that our results are not biased towards and not significantly influenced by a purely academic view.

We believe that, in comparison with focus groups and individual interviews, on-line surveys can be a highly valuable instrument in further investigations of this topic. There are two risks that can be mitigated by anonymous questionnaires:

1. In collaborations between academia and industry it is not unlikely that industrial participants in such projects are from the management, or senior engineers, or research engineers for several reasons not necessarily regularly connected to the operational teams. Such collaborations bear the risk that the sample gets biased towards these roles. With an on-line survey advertised on multiple channels, we are convinced to have mitigated such a bias.

2. For legal reasons, safety activities can be quite critical to talk about personally and in an open way. The authors’ experience and impression is that in personal interviews, practitioners tend to avoid talking loosely about their organisations and, where aggravating, to moderately generalise. Our impression from the respondents’ occasionally quite open comments leads us to believe that the risk of this bias is lower in anonymous surveys such as our questionnaire. Note that subjectivity has to be handled by other means in both questionnaires and interviews.

Leveson [49, p. 211] states that FMEA, with its limited applicability for safety analysis, is less frequently used as
a hazard analysis technique than FTA or ETA. As opposed to Leveson’s observation, our respondents most often state that they work(ed) with FMEA-based techniques in their safety activities (cf. Figure 8). One reason for this discrepancy could be that we only provided a small set of techniques as answer options to check the criterion Familiarity with Methods, particularly, ETA was not included. Assuming that many respondents are reluctant to add further techniques in the “Other” field, this might have led to a bias towards the specified answer options. Assuming that Leveson’s observation is drawn from US system safety cultures, this discrepancy could also have arisen from the circumstance that our sample is biased towards European safety cultures (cf. Geographical Regions in Section 4.2). While this issue limits the external validity of our exploratory study, we believe that the results for the questions qAdeqOfMthStd, qAppOfMeth, and qPosImpOfFMs and the hypotheses hInsufStds, hInsufMeth, hFMsImprSaf, hAdapt:AutoGTAero, hInsufMeth:EngDifSci, and hInsufMeth:AutoGTAero (relying on the constructs Familiarity with Methods, Means, and Adequacy) are not harmed by this issue.

The independence of most of the questions allows a per-question analysis. Particularly, the 59 partial responses might not affect any complete data points and thus were taken into consideration for the questions for which they delivered responses (cf. variation of \( N \) values). The relatively high number of registered views (565) might stem from users checking the questionnaire start page and concluding that they do not belong to the target group (Section 3.5): Diverse preconceptions of safety, diverse channel members, as well as short non-informative survey advertisements might have played a role. We believe, this issue has not led to a significant loss of relevant respondents or a participation of illegible respondents.

However, given that we expect the population of SPs to be 2 to 3 orders of magnitude larger than our sample (\( N = 91, N' = 124 \)), confident general conclusions cannot be drawn. For this, other sampling approaches such as the one employed by Manotas et al. [55] might be more appropriate, given proper multilateral backing and preparation. Their possibility to sample the population with the support of global software companies might be more effective than our approach based on volunteer and cluster-based sampling from several on-line discussion channels.

5.4.3. Reliability

To which extent would a repetition of the procedure in Section 3 with the same sample lead to the same results?

It is difficult to exactly repeat this survey in the short term because our advertisements covered many of the relevant on-line channels and we expect some of the respondents not willing to participate again within short-term or at all. This is a general problem for studies of this kind. Therefore, we suggest to 1. provide incentives, 2. pursue off-line channels as well, 3. repeat the study in the long term, and 4. extend the sampling period. For example, Mendez-Fernandez et al. [60] provide a longitudinal design supporting repeatability and hence the determination of reliability of the results.

5.5. Lessons Learned

Regarding the sample size (Section 4.2), we wished to get more responses against the background of the effort we had in reaching out to the population (Section 4.1). From the Unipark questionnaire view statistics, we saw that in some of the larger discussion forums, users seemed to appear noticeably reluctant to respond to our questionnaire. The return rates estimated in Section 4.1 can be considered low. In few discussion forums, our friendly, singular, and topic-related post of the questionnaire was even penalised by deleting the post or by loosing forum membership. Unfortunately, important non-commercial panels such as, for example, SoSciSurvey11 or SurveyCircle (Table 6) do not offer profiling facilities to focus on engineering professionals. In the case of no budget for incentives and for paying commercial panels, these circumstances make it very difficult for empirical (software) researchers to approximate a representative sample.

6. Conclusions and Future Work

We designed and conducted a questionnaire-based cross-sectional on-line survey of safety practitioners. Our objective was to investigate safety practice by asking practitioners about means they rely on, process factors influencing their work, and their role in the life cycle, and by checking several observations stemming from previous research.

6.1. Summary of Findings and Implications

Below, observations marked with + represent our aspirations when performing the study. Observations marked with − represent our apprehensions. Items labelled with

+ accommodate neutral observations.

We collected evidence in support of several hypotheses leading to the following observations:

+ Our respondents confirm that safety decision making is mostly based on expert opinion and experience from previous projects.

+ Safety practitioners think that for highly interconnected systems (e.g., systems of systems, connected transport systems), assurance of safety will have to rely on high assurance of IT security. Our experience suggests that the inverse relation is similarly strong.

11 See https://www.soscisurvey.de.
+ The survey participants believe that formal methods may have a positive impact on safety activities.

− Currently applied standards and practised methods are believed to be largely inadequate to cope with the assurance of technologies (e.g. adaptive control, machine learning) used for high automation and autonomy in upcoming system applications.

The last findings raise the question whether systems are safe enough and why this would be the case [10]?

Our analysis leads to further observations:

• Resources occasionally but not typically govern safety practitioners’ decisions and performance. The responses indicate that safety seems only rarely compromised by cost-benefit questions.

• Practitioners refrain from seeing safety as a special case of reliability. This stands in contrast with Leveson’s former observation that safety is pervasively confused with reliability [49].

− Safety practitioners think that many of their non-safety co-workers’ share at most medium appreciation of safety practitioners’ contributions to the life cycle.

− Respondents are indecisive on whether the conventional or ready-to-use methods they (could) apply scale sufficiently.

The last finding again motivates further analysis along the lines of [10]. If we are left unsure about whether means have become inadequate and, as found for safety RE in [12], if conventional approaches are dominant and we lack evidence for efficacy of novel research, how could safety research help practitioners?

In summary, we share the impression that empirical research in system safety is still in an early stage, on the one hand, offering many opportunities to perform cross-disciplinary studies and, on the other hand, bearing large risks of not exactly knowing to which extent safety practitioners are applying state of the art and able to do their best. This is a severe issue to be discussed in software and system safety research.

6.2. Future Work

We seek to extend our analysis by revisiting findings from the collected data set and not discussed in this work. Furthermore, we are going to identify and evaluate further hypotheses and ask more why- and how-questions.

Aspiring to the exploratory approach and grounded theory [53, p. 298], we can further engage with our survey participants using the focus group method [61], request for comments on our findings, and ask them for approaches to overcome the identified issues. Additionally, we re-shape our construct and focus on a smaller set of questions, for example, to investigate the applicability of formal methods.\(^\text{12}\)

Our research design can be extended towards the application of the goal question metric approach [62]: The results of the hypotheses analysis promotes the definition of goals of safety activities, the survey questions corresponding to the hypotheses can be refined, and process and product metrics be derived from the refined questions, for example, as already suggested and discussed in [63, 64]. Our study object includes SPs and, consequently, some of these metrics get measurable by questionnaires.

Our setting as well as our findings coin a good starting point for the design of a longitudinal study, offering possibilities to identify and validate causal relationships among the measured sub-constructs (Table 2).

Inspired by previous work [65] and by [24], it would be interesting to adapt our research design to support investigations of phenomena such as confirmation bias in practical safety arguments [39, 66].

Acknowledgments

The first author of this work is supported by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – GL 915/1-1. It is our pleasure to thank all survey participants for their valuable responses, and several practitioners, researchers, and students for acting as pilot run respondents and for providing us with initial feedback. Special thanks go to Mohammed Hussein and Dai Yang whose analyses yielded important preliminary findings for initiating this survey. We are indebted to Martin Wildmoser for attending the final interview for [4] with friendly support of the Validas AG\(^\text{13}\) in Munich and to further enthusiastic safety experts from various German industries for participation in the interviews for [5]. Technical University of Munich (TUM) and University of York have been excellent working environments. I would like to thank Manfred Broy for his senior advice and for providing the research infrastructure. Daniel Mendez-Fernandez deserves cordial gratitude for giving us Unipark advice and granting us access to this platform using the TUM Informatics faculty license. Finally, we would like to express our gratitude to the anonymous reviewers for helpful comments leading to significant improvements.

References

[1] C. Perrow, Normal Accidents: Living with High-Risk Technologies, Basic Books, NY, USA, 1984.

[2] J. Sorensen, Safety culture: a survey of the state-of-the-art, Reliability Engineering & System Safety 76 (2) (2002) 189 – 204. doi:10.1016/S0951-8320(02)00005-4.

[3] R. M. Choudhry, D. Fang, S. Mohamed, The nature of safety culture: A survey of the state-of-the-art, Safety Science 45 (10) (2007) 993 – 1012. doi:10.1016/j.ssci.2006.09.003.

\(^{12}\)The first author of this study has finished a follow-up survey on the use of formal methods, available via https://goo.gl/forms/FnKNOtTmJ3A6RekM2.

\(^{13}\)See http://www.validas.de.
M. Hussein, Current challenges of system safety practitioners: Qualitative analysis of on-line discussions, Master’s thesis, Technical University of Munich (2016).

D. Yang, Hazards from high system entropy: An exploratory analysis of case reports, Master’s thesis, Technical University of Munich (2016).

Health and Safety Executive, Out of Control, HSE Books, 2003.

F. P. Brooks, Jr., The Mythical Man Month: Essays on Software Engineering, 20th Edition, Addison-Wesley Longman, Amsterdam, 1995.

J. C. Knight, Safety critical systems: Challenges and directions, in: Proceedings of the 24th International Conference on Software Engineering, ICSE ’02, ACM, New York, NY, USA, 2002, pp. 547–50. doi:10.1145/581339.581406.

P. G. Neumann, Risks to the public, ACM SIGSOFT Software Engineering Notes 43 (2) (2018) 8–11. doi:10.1145/3203094.3203102.

J. McDermid, A. Rae, How did systems get so safe without adequate analysis methods?, in: 9th IET International Conference on System Safety and Cyber Security (2014), Institution of Engineering and Technology, 2014. doi:10.1049/cp.2014.0968.

L. J. Osterweil, Be gracious, ACM SIGSOFT Software Engineering Notes 43 (2) (2018) 4–6. doi:10.1145/3203094.3203100.

E. Bjarnason, P. Runeson, M. Borg, M. Unterkalmsteiner, E. Enderle, Requirements engineering for safety-critical systems: A systematic literature review, Information and Software Technology 75 (2016) 71–89. doi:10.1016/j.infsof.2016.04.002.

A. R. Nyokabi, Practical safety challenges: An online survey of safety practitioners demands, problems and expectations, Master’s thesis, Technical University of Munich (2017).

R. D. Alexander, A. J. Rae, M. Nicholson, Matching research goals and methods in system safety engineering, in: 5th IET International Conference on System Safety 2010, 2010, pp. 1–8. doi:10.1049/cp.2010.0822.

A. Rae, M. Nicholson, R. Alexander, The state of practice in system safety research evaluation, in: 5th IET International Conference on System Safety 2010, IET, 2010. doi:10.1049/cp.2010.0838.

R. Valerdi, H. L. Davidz, Empirical research in systems engineering: challenges and opportunities of a new frontier, Systems Engineering 12 (2) (2009) 169–181. doi:10.1002/sys.20117.

T. Dwyer, Industrial safety engineering—challenges of the future, Accident Analysis & Prevention 24 (3) (1992) 265 –273. doi:10.1016/0001-4575(92)90005-4.

B. Graaf, M. Lormans, H. Toetenel, Embedded software engineer: the state of the practice, IEEE Software 20 (6) (2003) 61–69. doi:10.1109/MS.2003.1241368.

J. Kasurinen, O. Taipale, K. Smolanov, Software test automation in practice: Empirical observations (2010). doi:10.1109/2010.620836.

J. Hachtly, A. Wassing, T. Kelly, C. Comar, P. Jones, Certifiable safe software-dependent systems: challenges and directions, in: Proceedings of the on Future of Software Engineering - FOSE 2014, ACM Press, 2014, pp. 182–200. doi:10.1145/2593882.2593895.

J. Chen, M. Goodrum, R. Metoyer, J. Cleland-Huang, How do practitioners perceive assurance cases in safety-critical software systems?, in: Proceedings of the 11th International Workshop on Cooperative and Human Aspects of Software Engineering - CHASE’18, ACM Press, 2018, pp. 57–60. doi:10.1145/3195836.3195838.

A. Cecarelli, N. Silva, Analysis of companies gaps in the application of standards for safety-critical software, in: F. Koornneef, B. A. Kitchenham, J. Singer, A. Forward, How software engineers use documentation: the state of the practice, IEEE Software 20 (6) (2003) 35–39. doi:10.1109/MS.2003.1241364.

A. Rae, R. Alexander, Forecasts or fortune-telling: When are expert judgements of safety risk valid?, Safety Science 99 (2017) 156–165. doi:10.1016/j.ssci.2017.02.018.

E. Bjarnason, P. Runeson, M. Borg, M. Unkermahler, E. Enderle, B. Regnell, G. Sabaliauskaite, A. Loconsole, T. Gorschek, R. Feldt, Challenges and practices in aligning requirements with verification and validation: a case study of six companies, Empirical Software Engineering 19 (6) (2013) 1809–1855. doi:10.1007/s10664-013-9263-y.

I. Dodd, I. Habli, Safety certification of airborne software: An empirical study, Reliability Engineering & System Safety 98 (1) (2012) 7–23. doi:10.1016/j.ress.2011.09.007.

E. Lim, N. Taksande, C. Seaman, A balancing act: What software practitioners have to say about technical debt, IEEE Software 29 (6) (2012) 22–27. doi:10.1109/MS.2012.130.

C. K. Sreh, Encyclopedia of Case Study Research, SAGE, 2010, Ch. “Exploratory Case Study”, pp. 372–3. doi:10.4135/9781412957397.

A. J. Rae, R. D. Alexander, Probative blindness and false assurance: An exploratory study, Reliability Engineering & System Safety 98 (1) (2012) 22–27. doi:10.1016/j.ress.2011.09.007.

T. Cant, System safety: Where next?, in: System Safety Conference incorporating the Cyber Security Conference 2013, 8th IET International, 2013, pp. 1–10. doi:10.1049/cp.2013.1706.

A. Knauss, J. Schroder, C. Berger, H. Eriksson, Paving the road for safety of automated vehicles: An empirical study on testing challenges, in: 2017 IEEE Intelligent Vehicles Symposium (IV), IEEE, 2017, pp. 1873–80. doi:10.1109/IVS.2017.799578.

J. Knight, Fundamentals of Dependable Computing for Software Engineers, Chapman & Hall/CRC Innovations in Software Engineering and Software Development, Chapman and Hall/CRC, 2012.

R. Bloomfield, P. Froume, B. Monahan, Formal methods in the production and assessment of safety critical software, Reliability Eng. & Sys. Safety 32 (1-2) (1991) 51–66. doi:10.1016/0951-8320(91)90047-B.

L. M. Barroca, J. A. McDermid, Formal methods: Use and relevance for the development of safety-critical systems, Comp. J. 35 (6) (1992) 579–99. doi:10.1093/comja/35.6.579.
A. Summary of All Responses

For validation purposes, the following tables present data summaries for all closed (q)uestions according to Table 4 and questions for classification according to Table 1. The “Option” column refers to the parts (if any) of multi-part questions. The “NAs” column signifies the number of invalid data points for each (part of a) question. The checksum (including invalid responses) of each row results in $N = 152$ responses. Rows with NAs = 0 result from parts (i.e., answer categories) added after content analysis of half-open questions (Section 3.6.1). The questions qEffRoleOnJob and qEffNotionOnJob are open and, hence, not accompanied by a corresponding table.

---

### qValueOfKnow

| Option / N | Value |
|------------|-------|
| low        | Very  |
| a / 96     | 1     |
| b / 96     | 1     |
| c / 95     | 1     |
| d / 95     | 1     |
| e / 96     | 9     |
| f / 96     | 1     |
| g / 97     | 1     |

| Impact |
|--------|
| Low    |
| Medium |
| High   |

Legend: a. Hazard list from previous projects, b. Case (accident, incident) reports, c. Inspection checklist, d. Expert opinions, e. Management recommendations, f. Co-workers’ recommendations, g. Safety-related project experience

### qImpOfConstr

| Option / N | Impact |
|------------|--------|
| a / 98     | 5      |
| b / 96     | 2      |
| c / 95     | 6      |
| d / 95     | 5      |
| e / 97     | 4      |
| f / 97     | 1      |
| g / 98     | 2      |

Legend: a. Budget cuts, b. Late or unclear choice of safety concepts, c. Postponed safety decisions, d. Schedule pressure, e. Erroneous hazard analyses, f. Vague safety standards, g. Inexperienced safety engineers

### qImpOfEco

| Option / N | Frequency |
|------------|----------|
| a / 98     | 36       |
| b / 96     | 48       |
| c / 95     | 9        |
| d / 95     | 6        |
| g / 98     | 53       |

Legend: a. Hazard list from previous projects, b. Case (accident, incident) reports, c. Inspection checklist, d. Expert opinions, e. Management recommendations, f. Co-workers’ recommendations, g. Safety-related project experience
### Adequacy

| Option | Do not know | Not adequate | Slightly adequate | Adequate | Very adequate | NA's |
|--------|-------------|--------------|-------------------|----------|---------------|------|
| a / 100| 32          | 31           | 26                | 11       | 0             | 52   |
| b / 101| 22          | 30           | 33                | 13       | 3             | 51   |
| c / 101| 26          | 31           | 27                | 16       | 1             | 51   |
| d / 101| 22          | 48           | 23                | 8        | 0             | 51   |
| e / 100| 36          | 7            | 25                | 27       | 5             | 52   |
| f / 101| 29          | 10           | 26                | 30       | 6             | 51   |
| g / 99 | 32          | 28           | 21                | 17       | 1             | 53   |

#### Legend:
- a. Security is prerequisite for safety
- b. Safety is prerequisite for security
- c. SPs depend on security practitioners
- d. Security practitioners depend on SPs
- e. Safety assurance requires security assurance
- f. Rare interaction in requirements stage
- g. Rare interaction in assurance stage
- h. Lack of collaboration is hazardous
- i. Lack of collaboration is inefficient
- j. Involvement in RE improves safety

### Agreement

| Option | Strongly Disagree | Neither agree nor disagree | Agree | NA's |
|--------|-----------------|--------------------------|-------|------|
| − / 102| 5               | 3                        | 28    | 11   | 50   |

#### Legend:
- a. Self-adaptive systems
- b. Highly automated and autonomous driving
- c. Distributed networked systems
- d. AI/ML-based applications
- e. Medical and healthcare applications
- f. Highly automated air traffic control
- g. Consumer or commercial drones

### Impact

| Option | No impact | Low impact | Medium impact | High impact | NA's |
|--------|-----------|------------|---------------|-------------|------|
| − / 62 | 4         | 3          | 15            | 23          | 17   | 90   |

#### Legend:
- a. Adapt skills to new technologies
- b. Study state-of-the-art safety principles
- c. Juniors learn from seniors
- d. Juniors learn from accident reports

### Agreement

| Option | Strongly Disagree | Neither agree nor disagree | Agree | Strongly agree | NA's |
|--------|-------------------|---------------------------|-------|---------------|------|
| a / 96 | 1                 | 1                        | 3     | 5             | 49   | 56   |
| b / 95 | 9                 | 14                       | 17    | 36            | 18   | 57   |
| c / 96 | 6                 | 4                        | 13    | 53            | 19   | 56   |
| d / 95 | 12                | 6                        | 18    | 21            | 32   | 57   |

#### Legend:
- a. Security is a cost factor
- b. Safety is a beneficial factor
- c. Safety is a necessity independent of cost
- d. Safety is a tedious mandated task
- e. Safety implies reliability

### Agreement

| Option | Strongly Disagree | Neither agree nor disagree | Agree | Strongly agree | NA's |
|--------|-------------------|---------------------------|-------|---------------|------|
| a / 97 | 1                 | 1                        | 7     | 44            | 44   | 55   |
| b / 97 | 3                 | 2                        | 3     | 40            | 42   | 55   |
| c / 97 | 2                 | 2                        | 16    | 44            | 31   | 55   |
| d / 97 | 2                 | 1                        | 7     | 41            | 44   | 55   |

#### Legend:
- a. Security is given high priority
- b. Management highly values safety
- c. SPs have declared authority
- d. Safety process is defined

### Agreed

| Option | Strongly Disagree | Neither agree nor disagree | Agree | Strongly agree | NA's |
|--------|-------------------|---------------------------|-------|---------------|------|
| a / 97 | 2                 | 57                       | 28    | 5             | 2    | 55   |
| b / 96 | 1                 | 11                       | 18    | 12            | 17   | 56   |
| c / 96 | 1                 | 4                        | 6     | 15            | 44   | 26   |
| d / 97 | 2                 | 17                       | 18    | 21            | 29   | 10   |
| e / 96 | 3                 | 25                       | 25    | 15            | 19   | 9    | 56   |

#### Legend:
- a. Lack of accidents reduces need for safety
- b. Accidents drive need for safety
- c. Accidents help SPs argue for safety
- d. Lack of accidents reduces need for safety
- e. Safety implies reliability

### Value

| Option | Very low | Low | Medium | High | Very high | NA's |
|--------|----------|-----|--------|------|-----------|------|
| a / 94 | 3        | 5   | 13     | 14   | 35        | 24   | 58   |
| b / 93 | 2        | 10  | 30     | 23   | 19        | 10   | 57   |
| c / 94 | 4        | 11  | 15     | 36   | 23        | 59   |
| d / 93 | 6        | 8   | 22     | 28   | 20        | 59   |
| e / 93 | 5        | 10  | 9      | 42   | 24        | 59   |

#### Legend:
- a. Adapt skills to new technologies
- b. Study state-of-the-art safety principles
- c. Juniors learn from seniors
- d. Juniors learn from accident reports

### Value

| Option | Very low | Low | Medium | High | Very high | NA's |
|--------|----------|-----|--------|------|-----------|------|
| a / 94 | 4        | 19  | 37     | 28   | 7         | 57   |

#### Legend:
- a. Adapt skills to new technologies
- b. Study state-of-the-art safety principles
- c. Juniors learn from seniors
- d. Juniors learn from accident reports

### Value

| Option | Very low | Low | Medium | High | Very high | NA's |
|--------|----------|-----|--------|------|-----------|------|
| a / 94 | 5        | 3   | 12     | 49   | 21        | 58   |
### qImpOfExp Agreement

| Option | Do not know | Strongly Disagree | Neither agree nor disagree | Strongly agree | NA's |
|--------|-------------|-------------------|---------------------------|----------------|------|
| a / 05 | 3           | 2                 | 6                         | 11             | 41   | 32   | 57   |
| b / 95 | 3           | 1                 | 2                         | 1              | 32   | 4    | 56   |

**Legend:** a. Senior SPs outperform junior SPs, b. Previous projects experience is beneficial, c. Adversarial thinking improves hazard analysis

### Educational Background

| Option | Checked | Unchecked | NA's |
|--------|---------|-----------|------|
| a / 124 | 49      | 75        | 28   |
| b / 124 | 13      | 111       | 28   |
| c / 124 | 23      | 101       | 28   |
| d / 124 | 31      | 93        | 28   |
| e / 124 | 22      | 102       | 28   |
| f / 124 | 12      | 112       | 28   |
| g / 124 | 6       | 118       | 28   |

**Legend:** a. Computer Science, b. Systems Engineering, e. Safety Science, f. Electrical and Electronics Engineering, g. Mechanical and Aerospace Engineering, h. Physics and Mathematics, k. Other Discipline

### Application Domains

| Option | Checked | Unchecked | NA's |
|--------|---------|-----------|------|
| a / 124 | 51      | 73        | 28   |
| b / 124 | 54      | 70        | 28   |
| c / 124 | 16      | 108       | 28   |
| d / 124 | 19      | 105       | 28   |
| e / 124 | 30      | 94        | 28   |
| f / 124 | 17      | 107       | 28   |
| g / 124 | 24      | 100       | 28   |
| h / 124 | 9       | 115       | 28   |
| j / 124 | 25      | 99        | 28   |
| l / 124 | 15      | 109       | 28   |
| m / 124 | 8       | 116       | 28   |
| o / 152 | 15      | 137       | 0    |
| p / 152 | 6       | 146       | 0    |

**Legend:** a. Automotive and Transport Systems, b. Aerospace Industry, c. IT Infrastructure and Networking, d. Power and Nuclear Industry, e. Industrial Processes and Plant Automation, f. Electronic Devices and Appliances, g. Healthcare Systems, h. Construction and Building Automation, j. Industrial Machinery, I. Naval Systems, m. Other Domain, o. Railway and Cablecar Systems, p. Military and Defense Systems

### Level of Experience

| Option / N | < 3 | 3 - 7 | 8 - 15 | 16 - 25 | > 25 | NA's |
|------------|-----|-------|--------|---------|------|------|
|            | 27  | 27    | 27     | 18      | 20   | 33   |

### Familiarity with Standards

| Option / N | Checked | Unchecked | NA's |
|------------|---------|-----------|------|
| a / 124    | 70      | 54        | 28   |
| b / 124    | 19      | 105       | 28   |
| c / 124    | 47      | 77        | 28   |
| d / 124    | 3       | 121       | 28   |
| e / 124    | 47      | 77        | 28   |
| f / 124    | 14      | 110       | 28   |
| i / 124    | 8       | 116       | 28   |
| k / 152    | 3       | 149       | 0    |
| l / 152    | 3       | 149       | 0    |
| m / 152    | 12      | 140       | 0    |
| n / 152    | 4       | 148       | 0    |
| o / 152    | 14      | 138       | 0    |

**Legend:** a. Generic, b. Machinery, c. Automotive and Transport, e. Agriculture, f. Aerospace and Avionics, h. Not Familiar, i. Other Standard, k. Nuclear and Other Energy, l. Medical Devices, m. Railway, n. Methodology and Tooling, o. Military and Defense

### Familiarity with Methods

| Option / N | Checked | Unchecked | NA's |
|------------|---------|-----------|------|
| a / 124    | 90      | 34        | 28   |
| b / 124    | 78      | 46        | 28   |
| c / 124    | 16      | 108       | 28   |
| d / 124    | 59      | 65        | 28   |
| e / 123    | 4       | 119       | 29   |
| f / 124    | 15      | 109       | 28   |
| g / 124    | 15      | 109       | 28   |
| i / 152    | 9       | 143       | 0    |
| j / 152    | 3       | 149       | 0    |
| k / 152    | 5       | 147       | 0    |
| l / 152    | 3       | 149       | 0    |
| m / 152    | 5       | 147       | 0    |
| n / 152    | 4       | 148       | 0    |
| o / 152    | 6       | 146       | 0    |
| p / 152    | 3       | 149       | 0    |

**Legend:** a. Failure Mode Effects, b. Fault Trees, c. STAMP-based Methods, d. Hazard Operability, e. Security Threats, f. Not Familiar, g. Other Concept/Method, i. Functional and Design Risk, j. Probabilistic Risk, k. Assurance Cases, l. Root Causes, m. Event Trees, n. Automated V&V, o. Dependencies and Interactions, p. Bidirectional Methods

### Native Languages

| Option / N | Checked | Unchecked | NA's |
|------------|---------|-----------|------|
| a / 124    | 54      | 70        | 28   |
| b / 124    | 38      | 86        | 28   |
| c / 124    | 5       | 119       | 28   |
| d / 124    | 6       | 118       | 28   |
| e / 124    | 2       | 122       | 28   |
| g / 152    | 3       | 149       | 0    |
| h / 152    | 3       | 149       | 0    |
| i / 152    | 14      | 138       | 0    |

**Legend:** a. English, b. German, c. Italian, d. French, e. Chinese, g. Portuguese, h. Swedish, i. Other language

### Working Languages

| Option / N | Checked | Unchecked | NA's |
|------------|---------|-----------|------|
| a / 124    | 113     | 11        | 28   |
| b / 124    | 41      | 83        | 28   |
| c / 124    | 5       | 119       | 28   |
| d / 124    | 10      | 114       | 28   |
| e / 124    | 12      | 112       | 28   |

**Legend:** a. English, b. German, c. Italian, d. French, e. Other language
Table B.9: List of abbreviations used in this article

| Abbreviation | Description |
|--------------|-------------|
| AC | Acceptance Criterion |
| CAST | Causal Analysis using System Theory |
| CCA | Common Cause Analysis |
| CIA | Change Impact Analysis |
| CMA | Common Mode Analysis |
| DV | Dependent Variable |
| ETA | Event Tree Analysis |
| FFA | Functional Failure Analysis |
| FHA | Functional Hazard Analysis |
| FM | Formal Method |
| FMEA | Failure Mode Effects Analysis |
| FMEDA | Failure Mode, Effects, and Diagnostic Analysis |
| FRAM | Functional Resonance Analysis Method |
| FMECA | Failure Mode, Effects, and Criticality Analysis |
| HRA | Health Risk Assessment |
| IV | Independent Variable |
| MC | Multiple Choice |
| NA | Not Available |
| OSHA | Operation & Support Hazard Analysis |
| PHA | Preliminary Hazard Analysis |
| PHL | Preliminary Hazard List |
| RCA | Root Cause Analysis |
| RE | Requirements Engineering |
| RQ | Research Question |
| SACM | Structured Assurance Case Meta-Model |
| SCRA | Supply Chain Risk Assessment |
| SE | Software Engineering |
| SHA | System Hazard Analysis |
| SHARD | Software Hazard Analysis and Resolution in Design |
| SP | Safety Practitioner |
| SPP | Safety Practice and its Practitioners |
| STAMP | System-Theoretic Accident Model & Processes |
| STPA | System-Theoretic Process Analysis |
| STRIDE | Spoofing, Tampering, Repudiation, Information disclosure, Denial of service, Elevation of privilege |
| SSHA | System Safety (or Sub-System) Hazard Analysis |
| WBA | Why-Because Analysis |
| ZHA | Zonal Hazard Analysis |
| ZSA | Zonal Safety Analysis |

### Safety-related Roles

| Option / N | Checked | Unchecked | NA's |
|------------|---------|-----------|------|
| a / 124    | 62      | 62        | 28   |
| e / 124    | 29      | 95        | 28   |
| f / 124    | 9       | 115       | 28   |
| g / 124    | 39      | 85        | 28   |
| h / 124    | 23      | 101       | 28   |
| j / 124    | 11      | 113       | 28   |
| k / 124    | 3       | 121       | 28   |
| m / 66     | 12      | 54        | 86   |
| n / 124    | 4       | 120       | 28   |
| q / 152    | 3       | 149       | 0    |

Legend: a. Safety Practitioner, e. Software Practitioner, f. Electrical or Electronics Practitioner, g. Systems Practitioner, h. Requirements Practitioner, j. Health and Safety Practitioner, k. IT Security Practitioner, m. Reliability Practitioner, n. Other Role, q. V and V Practitioner

| Option / N | Checked | Unchecked | NA's |
|------------|---------|-----------|------|
| b / 124    | 26      | 98        | 28   |
| c / 124    | 19      | 105       | 28   |
| d / 124    | 5       | 119       | 28   |
| 1 / 76     | 13      | 63        | 76   |
| p / 152    | 12      | 140       | 0    |

Legend: b. Researcher in Industry, c. Researcher in Academia, d. Assistant, Trainee, or Junior, l. Practitioner with PhD degree, p. Consultant / Assessor

B. List of Abbreviations

See Table B.9.