

From Theory to Practice: Unraveling Epistemological Threads in Human Reliability Assessment

Pedro Baziuk^{1,2,3}, Juan Leiva Butti^{1,2}, Franco Macello², Oliverio Gabrielli², Ricardo Greco³

¹Faculty of Economic Sciences, National University of Cuyo, Mendoza, Argentina

²Faculty of Engineering, National University of Cuyo, Mendoza, Argentina

³ESEADE Institute, Buenos Aires, Argentina

Email: pedro.baziuk@fce.uncu.edu.ar

How to cite this paper: Baziuk, P., Butti, J. L., Macello, F., Gabrielli, O., & Greco, R. (2024). From Theory to Practice: Unraveling Epistemological Threads in Human Reliability Assessment. *Open Journal of Philosophy*, 14, 315-339.
<https://doi.org/10.4236/ojpp.2024.142021>

Received: February 1, 2024

Accepted: April 25, 2024

Published: April 28, 2024

Copyright © 2024 by author(s) and Scientific Research Publishing Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

Abstract

This comprehensive review explores the evolving epistemological foundations of Human Reliability Analysis (HRA), tracing its transition from a practical, applied field to a scientific discipline. The paper critically examines the philosophical assumptions that underlie prominent HRA models, emphasizing their impact on both scientific validity and practical application. Organized into five sections, the paper begins by articulating its primary aim in Section 1: to uncover and scrutinize the philosophical assumptions within HRA models. Sections 2.2 and 2.3 delve into reductionist and holistic paradigms. Section 2.4 focuses on the shift toward holistic theories, emphasizing the integration of human, technical, and organizational elements in modern socio-technical systems. Section 3 concludes by examining systemic approaches, capturing the paradigm shift from reductionism to a broader, interdependent view of accidents and reliability. Section 4 sheds light on the contemporary HRA models, emphasizing the prevalence of systemism in recent research. Notable models, such as SoTe-RiA and FRAM, undergo scrutiny for their systemic approach, revealing interrelationships crucial for understanding accidents in complex operations. In Section 5, the paper explores the practical implications of the epistemological insights gained, emphasizing their potential impact on the development and enhancement of HRA models. The conclusion advocates for a paradigm shift in HRA that involves rethinking the epistemological bases of models. It underscores the need for ongoing philosophical scrutiny to drive the discipline's progress, ultimately enhancing our understanding, prediction, and improvement of human performance and system safety.

Keywords

Human Reliability Analysis, Epistemological Foundations, Systemic Approaches

1. Introduction

Human Reliability Analysis (HRA) stands at the nexus of critical importance within the broader domain of safety science and engineering. As our reliance on complex systems continues to grow, understanding and mitigating human errors become paramount for ensuring the safety and reliability of socio-technical systems. HRA, which originated from the engineering need to measure and predict human errors in high-stakes industrial settings, has evolved into a field that bridges the realms of applied science and theoretical inquiry.

The motivation behind focusing on HRA in this comprehensive review is rooted in the transformative journey the discipline has undergone—from its pragmatic, applied origins to its current aspiration to establish itself as a scientific discipline. This paper seeks to delve into the philosophical underpinnings of HRA models, dissecting the implicit assumptions that guide their development. By scrutinizing the epistemological foundations, we aim to unravel the intricate relationship between theoretical perspectives and practical applications within HRA.

Understanding the underlying reasons for this focus on HRA is crucial, as it not only shapes the direction of our exploration but also informs the broader discourse on the intersection of human factors, system safety, and scientific rigor. Through this review, we aim to shed light on the significance of epistemological analysis in advancing HRA, contributing to ongoing conversations about the robustness, reliability, and theoretical coherence of models designed to enhance safety in complex operational environments.

Human Reliability Analysis (HRA), as defined by [Munger et al, \(1962\)](#), emerged as a field within engineering driven by the imperative to quantify and predict human errors in critical industrial settings. Rooted in the pragmatic ethos of efficiency and measurement associated with engineering studies, HRA has evolved from a practical, applied discipline to one that seeks scientific legitimacy.

The definition and conceptualization of HRA have evolved over time, shaped by the interdisciplinary nature of the field, drawing insights from anthropology, psychology, and sociology. This comprehensive review delves into the epistemological foundations of HRA, scrutinizing the implicit philosophical assumptions that underpin prominent models and methodologies.

HRA models serve three primary purposes: (1) quantifying human error probabilities (HEP), (2) describing and modeling human errors within industrial environments, and (3) enhancing the resilience of human-machine systems. Ensuring the theoretical plausibility of HRA models is crucial, as emphasized by [Kirwan \(2017\)](#), requiring support from experimental data and empirical validation, as advocated by [Mosleh and Chang \(2004\)](#).

The core elements of most HRA models encompass: (1) a human error taxonomy which classifies human error types along with their distinctive characteristics; (2) a methodological module for quantifying human errors, considering the natural inclination to err, and incorporating environmental, organizational,

and personal influences; and (3) presenting results in the form of guides or recommendations to mitigate human errors, thereby increasing human reliability and system resilience.

From an epistemological standpoint, the choice between holistic and reductionist approaches becomes pivotal when analyzing the human factor. According to Machado Susseret (2005), both are significant research programs in contemporary science. However, Bunge (2000) contends that methodological individualism, viewed as rational reductionism, proves ineffective both as a methodology and in its alternative forms.

Ontological reductionism posits a robust thesis: in deeper analysis, everything is either an individual or a group of individuals (Udehn, 2002). This perspective asserts that reality can be reduced to a minimal number of entities or basic components. For instance, in biology, organisms are exhaustively composed of the same components as inorganic matter (Ayala, 1987). The implication is that no totalities possess their own properties, be they systemic or emergent.

Contrarily, systemic thesis, rooted in holism, asserts that everything in the world is connected, either directly or indirectly (Bunge, 2000). Individualists view elements as mutually independent, as reflected in many HRA models, while holists prioritize the analysis of the whole and argue that relationships precede the related elements. Neither individualist ontology nor holism can entirely and deeply describe many phenomena, a point that will be explored further in the context of human error. The systemic approach contends that everything is a system or a component of one or more systems with emergent or systemic properties, requiring analysis of the system's composition, structure, and environment.

Machado Susseret (2005) suggests that both methodological individualism and the holistic approach have gained prominence as research programs, given their ability to address a wide range of phenomena and the elegance of their theories, endowed with strong heuristic and hermeneutical power. However, Bunge concludes that systemic thinking is the correct alternative to any form of individualism and holism.

An epistemological approach is essential, even in engineering sciences, with significant ontological, gnosiological, and methodological consequences, especially in the study of human reliability. While practical disciplines often overlook the philosophical assumptions implicit in advocated methodologies, these assumptions carry practical ramifications, as noted by Abraha and Liyanage (2012) in their review of theoretical foundations for risk minimal operations in complex sociotechnical systems, emphasizing the role of human error.

This paper seeks to provide a comprehensive exploration of the epistemological underpinnings in Human Reliability Analysis (HRA) models, aiming to contribute to the ongoing maturation of HRA as a scientific discipline. The central objective is to scrutinize the philosophical assumptions implicitly or explicitly embedded in prominent HRA models, shedding light on their implications for understanding and predicting human performance and system safety. The paper

is structured into distinct sections, each focusing on a key epistemological approach—reductionism, holism, and systemism. Through an extensive review, the text unveils the intricate interplay between these epistemological perspectives and the evolution of HRA, offering insights into how each approach has influenced model development and contributed to the broader discourse in human reliability studies. In essence, this paper endeavors to foster a deeper understanding of the philosophical foundations of HRA models, providing a critical examination that may catalyze future advancements in the field.

The examination of these philosophical positions within the realm of Human Reliability Analysis (HRA) has been relatively limited, primarily due to the practical orientation of HRA toward the development of useful models and methods rather than active participation in philosophical discourse. For instance, seminal HRA models like the Human Error Assessment and Reduction Technique (HEART) or the Technique for Human Error Rate Prediction (THERP) have historically focused on developing practical tools for industry application rather than engaging in explicit philosophical discourse. These widely utilized models, while invaluable for practical purposes, often lack explicit articulation of their underlying philosophical assumptions. The emphasis on crafting actionable methodologies has understandably steered the HRA community toward pragmatic outcomes, inadvertently contributing to a scarcity of philosophical examinations within the field. By delving into specific instances where HRA models prioritize practical utility over philosophical discourse, this study aims to shed light on the prevailing trend and underscore the need for a more balanced integration of theoretical considerations within the HRA discourse.

However, this review holds significant potential for several reasons: Firstly, philosophical preconceptions inherently shape scientific approaches, making it imperative to scrutinize these aspects in HRA. Addressing such fundamental issues not only enhances scientific rigor but also strengthens methodological validations, fortifies model robustness, establishes a conceptual framework for the discipline's scientific standing, and propels future progress. Secondly, for newcomers to the academic and scientific study of human reliability and human factors, this paper serves as an introduction to the field, addressing questions such as the origins of models and the gnosiological pathways that guided researchers in developing specific models. Thirdly, it offers an updated perspective on HRA models, concepts, and trends, incorporating a review of scientific HRA articles up to 2022. Lastly, the review may serve as a guiding light for researchers, illuminating potential avenues for new developments in the field.

2. Critical Review of Epistemological Paradigms in HRA

2.1. Material and Method

The primary objective of this study is to meticulously assess the epistemological foundations within Human Reliability Analysis (HRA) models. The focus is on uncovering and critically examining the philosophical assumptions that signifi-

cantly influence the assessment of HRA models. This assessment aims to shed light on the intricate interplay between philosophical underpinnings, scientific validity, and the practical application of HRA methodologies. It operates under the assumption that these models initially adopted a reductionist approach, eventually transitioning through a holistic stage to embrace a systemic perspective. While HRA has traditionally focused on delivering practical models and methods, philosophical underpinnings play a pivotal role in shaping any scientific endeavor.

This exploration of HRA methods from an epistemological standpoint provides numerous benefits: (1) Uncovering the philosophical preconceptions inherent in these models is critical for a deeper scientific understanding. It facilitates rigorous scrutiny, enhances methodological validations, fortifies model robustness, establishes a conceptual framework for the scientific positioning of the discipline, and contributes to future progress and trends; (2) For newcomers in the academic study of human reliability and human factors, this paper serves as an introduction, addressing questions such as the origins of models and the gnosiological pathways guiding researchers in model development. It delves into the philosophical assumptions that underpin human reliability approaches; (3) It offers an updated perspective on HRA models, concepts, and trends, encompassing scientific articles in the field up to 2022; (4) For HRA researchers, it acts as a source of inspiration, providing ideas and prompting insights for the revision of existing models or the creation of new ones.

Scientific inquiry is a multifaceted process guided by various epistemological approaches, each offering distinct perspectives on how knowledge is generated and understood. Reductionism, holism, and systemic approaches represent three fundamental paradigms that shape the foundations of scientific thinking and methodology.

Reductionism posits that complex phenomena can be understood by breaking them down into simpler, more fundamental components. It seeks to analyze and explain intricate systems through the examination of their individual parts. Reductionist approaches are characterized by a focus on isolating specific variables, often in controlled settings, to uncover underlying principles. While reductionism has been immensely successful in certain scientific domains, it has been criticized for oversimplifying complex interactions and neglecting emergent properties that arise from the collective behavior of interconnected components.

Anderson (Anderson, 1972) argues that the ability to reduce everything to simple fundamental laws does not imply reconstructing the universe based on those laws. This implies two difficulties, one of scale and another of complexity. The first is due to the fact that scale changes involve properties changes that are not always predictable. And the second difficulty concerns the fact that entirely new properties appear when complexity is increased, which is called emergent properties. Following Anderson's (Anderson, 1972) argument, psychology is not applied neuroscience; neuroscience is not applied molecular biology; and molecular biology, applied chemistry. The whole is not the sum of its parts, but above

all, very different from those.

In contrast, holism emphasizes the interconnectedness of elements within a system and contends that the whole is more than the sum of its parts. Holistic approaches advocate for studying phenomena in their entirety, considering the intricate relationships and interactions that contribute to their emergence. Holism recognizes the significance of context and the dynamic nature of systems, encouraging a more integrative and comprehensive understanding. However, critics argue that holism might risk overlooking crucial details and mechanisms inherent in reductionist analyses.

Systemic approaches combine elements of both reductionism and holism, acknowledging the importance of individual components while also considering their collective interactions. These approaches view systems as integrated entities with properties and behaviors that cannot be fully grasped by examining their isolated parts. Systemic thinking permeates various scientific disciplines, recognizing the complexity of real-world phenomena and the need for interdisciplinary perspectives. It often involves modeling the interactions between components to understand the system's emergent properties and behaviors.

This paper explores the epistemological foundations of reductionist, holist, and systemic approaches, aiming to elucidate how these paradigms influence scientific inquiry across diverse fields. By critically examining their implications and applications, we seek to contribute to a deeper understanding of the philosophical underpinnings that shape the trajectory of scientific research.

A comprehensive literature review methodology is employed in this paper to explore the epistemological foundations of reductionist, holist, and systemic approaches. By critically examining existing research, the aim is to elucidate how these paradigms influence scientific inquiry across diverse fields. The analysis encompasses philosophical implications, historical contexts, and practical applications, providing a nuanced understanding of the philosophical underpinnings shaping the trajectory of scientific research.

This article primarily constitutes a literature review, drawing insights from research articles, manuals, and available literature relevant to the discussed topics. All referenced documents were accessed from prestigious databases, including ISI-Web of Science, Science Direct, Springer Link, Informaworld, Engineering Village, Emerald, and IEEE Xplore, accessible through a subscription from the Argentinean National Council for Scientific and Technical Research (CONICET). The analysis focuses on papers published before 2022, encompassing a timeframe of approximately 30 years, with the earliest papers in the dataset dating back to the 1980s. However, the core of the study is rooted in the period from 1990 to 2022.

The principal objective is to investigate how “holism,” “reductionism,” and “system theory” have exerted ontological, gnosiological, and methodological influences on the conception, definition, development, and progression of HRA models and methods. The underlying hypothesis posits that HRA researchers, either explicitly or implicitly, make epistemological and philosophical assumptions,

derived from abductive reasoning informed by theoretical reviews and practical applications of HRA techniques. To validate this hypothesis, systematic searches for evidence of these assumptions were conducted within scientific databases.

The bibliometric analysis yielded an extensive collection of scientific publications (almost 10,000 articles), wherein terms like “holism” or “holistic” (21%), “reductionism” or “atomistic” (31%), and “systemic approach” or “system approach” (48%) were associated with “human reliability analysis” (263 articles), “human error” (3095 articles), “safety science” (1030 articles), and “risk analysis” (5410 articles). Additionally, terms such as “complexity” or “complexities” and “emergence” or “emergent” were explored concerning “human reliability analysis” and “human error.”

A research trend analysis indicates a growing interest in these topics linked with human reliability, human error, safety science, and risk analysis (48% of the articles were published after 2012), with a notable emphasis on the “systemic approach” (articles featuring the “systemic approach” increased their share in publications from 48% to 55% in the last five years).

The search results are organized by academic databases through ranking algorithms, typically offering options to sort results by date, title or journal alphabetically, citation counts, downloads, or relevance.

Relevance ranking involves multiple criteria, often indicating that a document is more relevant if a search term occurs frequently within it. For instance, in Web of Science, results are ranked based on the overlap between search terms and terms in the articles. In Scopus, the relevance rank considers the relative frequency and location of search terms in the article. In IEEE Xplore, the ranking is based on how well the result matches the search query, as determined by IEEE Xplore. Google Scholar, one of the few academic search engines combining various approaches in a single algorithm, weighs the full text of each document, where it was published, who wrote it, and how often and recently it has been cited in other scholarly literature .

Locating significant articles for this philosophical analysis of HRA involved multiple searches and a meticulous selection process due to the substantial number of articles retrieved.

In the initial selection phase, the 20 most relevant articles (according to search engine ranking algorithms), along with the 10 most cited and 10 most recent, were chosen. Following a preliminary screening analysis, articles using terms in different contexts or merely mentioning them were excluded. Subsequently, in a second step, an in-depth search was conducted, employing boolean operators like AND, OR, and NOT to identify significant articles not covered in the initial selection.

Furthermore, an exploration and analysis of publications related to widely used Human Reliability Analysis (HRA) methods were conducted to identify connections with the analyzed epistemological terms. Following the methodology outlined by Bell and Holroyd (Bell & Holroyd, 2009), certain methods, recog-

nized for being among the most cited, reviewed, criticized, and publicly available for implementation, were found to have associations with the examined epistemological terms.

The second hypothesis posits an epistemological evolutionary process from first to second-generation HRA models, suggesting that first-generation models predominantly embraced reductionism, while second-generation models leaned more towards holism. Additionally, recent HRA developments, often referred to as the third generation, exhibit a significant systemic tendency. To scrutinize this hypothesis, an investigation into the most notable characteristics of HRA generations was conducted to identify epistemological evidence.

The results are presented in two sections: Section 2.2 provides a summary of evidence supporting reductionism in the first generation, 2.3 holism in the second generation, and 2.4 epistemological assumptions in nine HRA methods. In Section 3, assumptions related to systemic approaches, including emergency and complexity theories, in recent HRA methods are detailed. Finally, Section 4 delves into the practical consequences derived from these epistemological considerations.

2.2. Decoding First Generation HRA: A Deeper Look into Reductionist Paradigms

The inception of Human Reliability Analysis (HRA) methods dates back to the 1960s, with significant advancements occurring in the mid-80s, particularly in terms of evaluating human factors' propensity to fail. These techniques can be broadly categorized into two generations: first and second. While dynamic HRA methods of the third generation are currently under research (Di Pasquale et al., 2013), this section delves into the epistemological underpinnings of the initial phase, specifically focusing on reductionism in First Generation HRA.

The study of human reliability has historically embraced both reductionist and holistic approaches (Brewer, 2006; Boring et al., 2005; Bodsberg, 1993). However, recent model developments and criticisms of previous generations suggest a shift towards a systemic approach. White (1995) underscore reductionist and holistic features in risk management theories, implying that prevalent risk management approaches lean towards reductionism.

The analysis of Reductionist Evidences in First Generation HRA (Table 1) reveals the hierarchical breakdown of tasks and the emphasis on observable characteristics, aligning with the methodology employed for this critical review.

Reduction, as elucidated by Bunge (2003), involves identifying or including objects or concepts within others, a concept with ontological, gnosiological, and methodological dimensions. Notably, while reductionism is based on reduction, utilizing reduction does not necessarily imply reductionism. For instance, the hypothesis of mind-brain identity posits that mental processes are reducible to brain processes—an ontological reduction—yet it doesn't advocate total reductionism in psychology to neurophysiology.

Table 1. Reductionist evidences in first generation HRA.

Main features	Reduction approach evidences
Binary representation of human actions (failure/success)	An ontological reduction is made when human actions are reduced to two types of actions.
Dichotomy of errors (omission and commission)	Tasks are broken into procedural steps through hierarchical task analyses (HTA) (Dekker, 2005).
Attention on human actions' observable characteristics and their classification (human error taxonomy)	Taxonomy approach comes from biology and tries to classify entities. Methodological reductionism assimilates a problem to a known conceptual scheme and is expected to make intelligible the basic structure of the present problem.
Low attention to cognitive actions (lack of a cognitive model)	Human actions are explained at their basic level (methodological reductionism), observed behavior, or physical actions. Reduction does not imply, in this case, an appeal to other scientific concepts but, rather, to choose a few ideas and try to explain with them all reality or a vast sector of it. "Mind is understood as a box-like construction with a mechanistic trade-in internal representation" (Abraha & Liyanage, 2012).
Emphasis on quantifying errors by statistical methods	HRA is described in terms of hardware reliability, i.e., human behavior is addressed in terms of mechanical components (mechanistic reductionist assumptions (French, Bedford, Pollard, & Soane, 2011) (Sornette, Maillart, & Kröger, 2013). Statistical methods are based on binary logic where membership degrees or degrees of truth are not allowed. Ontological reductionism opposes any form of dualism (Block & Stalnaker, 1999).
Indirect treatment of context with PSF	PSFs are basic characteristics (entities) of context. Most theories are based on implicit functions that relate PSFs with error probabilities. Gnosiological reductionism argues that the best strategy is to search for explanations in terms of the basic constituent entities of objects or processes under study.

Methodological reductionism posits that studying phenomena at the lowest levels of complexity is the optimal research strategy, proposing that constructs are explained from individual concepts. Bunge (2000) argues that predicates precede types, particularly in the case of membership sets defined by set theory axioms. Many HRA taxonomies seem to align with a moderate individualistic approach, where elements together define predicates, as seen in PSF sets.

First-generation models, exemplified by THERP (Technique for Human Error-Rate Prediction), are characterized by binary representation of human actions, dichotomy of errors, attention to human action phenomenology, minimal focus on cognitive actions, reliance on statistical methods for error quantification, and indirect treatment of context (Kim, 2001). THERP, based on event tree analysis, exemplifies a reductionist approach where emergent properties arising from the whole system may not be adequately recognized (White, 1995).

While THERP, along with parallel approaches like HCR (Human Cognition Reliability), adopts a cognitive model of human behavior, known as the skill-rule-knowledge model (SRK), this model reflects an individualistic approach by decomposing tasks into subtasks. However, a common critique of first-generation models is their failure to incorporate environmental, organizational, and other relevant factors—a limitation attributed to a lack of a holistic perspective (French et al., 2011).

The reductionist nature of first-generation models becomes evident in their difficulty to capture the underlying causes of human error in dynamic and complex actions. Moreover, they often overlook interactive combinations of equipment failures or common cause failures, a limitation addressed by more comprehensive models like “Failure Mode and Effects Analysis” (FMEA) (White, 1995). Probabilistic methods also face challenges in dealing with the variability, uncertainty, and incomplete knowledge inherent in many domains (Gregoriades & Sutcliffe, 2008).

The major drawbacks of probabilistic methods include a lack of reliable information, insufficient criteria for selecting Performance Shaping Factors (PSFs), limitations in assessing cognitive behavior, and treating human errors as phenomena without sufficient attention to their causes (Richei et al., 2001). The need for moderate reductionism is acknowledged, recognizing the heuristic value of nonreductionistic positions (Ayala, 1987).

Despite its justifications, reductionist HRA approaches may not be well-suited for complex socio-technical systems, where behavior emerges holistically, is highly sensitive to small input changes, and can only be partially described subjectively (Zio, 2009). Traditional linear cause-and-effect models may not effectively capture the dynamics of current organizational structures, prompting the introduction of complexity theory concepts in HRA. This shift calls for holistic approaches, especially when considering team effects and system dynamism (French et al., 2011; Stanton, 2016).

2.3. Navigating the Cognitive Landscape: Unveiling Second-Generation HRA Paradigms

The early 1990s witnessed a surge in research and development activities aimed at enhancing Human Reliability Analysis (HRA) methods globally. These endeavors resulted in significant strides in first-generation methods and the emergence of innovative techniques, marking the advent of second-generation HRA methods. Initially obscured, these methods were characterized by their conceptual aspirations, a departure from the primarily behavioral focus of first-generation HRA methods (Mosleh & Chang, 2007). The second-generation HRA methods sought to unravel the intricacies of cognitive aspects, delve into the causes of errors rather than their frequency, and explore the interaction and interdependence of factors, including Performance Shaping Factors (PSFs).

Epistemologically, this shift from reductionism to holism was deemed necessary to effectively model the complexity inherent in human and organizational

risk (Abraha & Liyanage, 2012). This complexity, as emphasized in the paper authored by Abraha and Liyanage (2012), cannot be fully understood or explained by the summation of individual components. Modern or second-generation models, as outlined by Cacciabue (2000), exhibit distinctive features such as a consideration of cognitive and organizational factors, a reference to cognitive and/or group/organization models, and the necessity of being conducted by a team of experts.

In the second-generation models, the emphasis is placed on understanding error causes rather than error frequency, with a focus on qualitative aspects, interaction, and interdependence of factors (Di Pasquale et al., 2013). Conceptual constructs like safety culture and root cause analysis take a more holistic view of risk, aiming to describe failures and explain their occurrence (White, 1995). Table II succinctly summarizes the main features of second-generation models and their holistic evidences, developed by the author of this paper.

Advancing towards a more holistic approach, second-generation models challenge the conventional idea of human error taxonomy. Latorella & Prabhu (2000) underscore the need for a holistic approach to classify human errors, asserting that errors should not be considered in isolation as a distinct class of behaviors. Introducing a new category of error, “cognitive error,” these models acknowledge the impact of technological advances on reducing physical activity-related errors while amplifying the consequences of reasoning or cognitive errors deeply rooted in socio-technological contexts (Cacciabue, 2000).

Cognitive errors, typically associated with a human behavior model, prompt a departure from linear information processing models to cyclical models. This shift aligns with a broader move from reductionist to holistic thinking, treating human behavior as a holistic interaction with the environment, incorporating the role of decision-making processes.

In the table dedicated to Holistic Evidences in Second Generation HRA (Table 2), main points include the consideration of cognitive and organizational factors, the dichotomy in cognition models (microcognition and macrocognition), reliance on expert judgment, emphasis on error causes, and the incorporation of conceptual constructs. This table, developed by the author of this paper, underscores the holistic paradigm that guides the discussion of second-generation HRA models.

2.4. Mapping Epistemological Approaches in Human Reliability Analysis Methodologies

The Exploration of Epistemological Approaches in Human Reliability Assessment (HRA) Methods is vital for understanding the intricacies of different methodologies. In Table 3, a comprehensive review of principal HRA methods and their epistemological characteristics is presented, offering insights into the reductionist and holistic aspects inherent in each approach.

Addressing the nuances in reductionist and holistic interpretations is crucial. For instance, the term “black box” HRA models, though often perceived as holistic due to their lack of detailed human behavior modeling, may contradict this

view according to Pyy (2000). Conversely, decomposed HRA modeling approaches, such as THERP and ATHEANA, may be labeled reductionist by emphasizing internal failure mechanisms (Pyy, 2000). Furthermore, the use of Probabilistic Safety Factors (PSF) differs, with some methods explicitly incorporating PSFs (e.g., THERP, ATHEANA, SLIM), while holistic expert judgment methods do so implicitly.

The concept of holism is occasionally synonymous with a systemic approach or complex thought. While systems thinking adopt a holistic approach, analytical methods in risk assessment tend to lean towards reductionism (White, 1995). In contemporary scenarios, where complex systems lack well-defined boundaries, set targets, and historical data, traditional reliability engineering methodologies face limitations (Abraha & Liyanage, 2012; Zio, 2009). Both first and second-generation HRA methods struggle with emergent and complex factors in new settings, hindering the exploration of underlying error causality (Abraha & Liyanage, 2012).

Table 2. Holistic evidences in second generation HRA.

Main features	Reduction approach evidences
Cognitive and Organizational Factors	Holistic models find application in complex dynamic systems, allowing the emergence of relationships among technical, human, and organizational factors.
Cognitive Model and/or Group/Organization Model	The dichotomy between reductionism and holism is evident in cognition models, namely microcognition and macrocognition. While microcognition dissects individual components like memory and attention, macrocognition focuses on cognitive functions such as decision-making, emphasizing external validity and regularity across conditions (Farrington-Darby & Wilson, 2006).
Based on Expert Judgment	Experts perceive events and circumstances as irreducible wholes (Boring, Gertman, Joe, & Marble, 2005). The formalism of structured expert elicitation addresses the highly complex and interdependent conditions involved in evaluating human reliability, enabling analysts to consider a realistic set of evidence holistically (Forester et al., 2004).
Focus on Error Causes	The “etiology of hazards” in terms of human reliability seeks the origin of the chain of causes leading to an accident. While reductionist approaches may suffice in blaming human error for an accident, holistic approaches consider human error as the starting point for investigation. Once identified, the causes of the error are systematically determined, investigated, and mitigated (Abraha & Liyanage, 2015; Abraha & Liyanage, 2012).
Conceptual Constructs	In contrast to classical inductivism dominating first-generation HRA, second-generation models are grounded in conceptual constructs aiming to elucidate reality. These constructs, such as situation awareness, complacency, and effective crew resource management, serve as holistic causal factors and aid in drawing conclusions about accidents. They bridge observable behavior with non-observable structures (Dekker & Hollnagel, 2004).

Table 3. Epistemology approaches in HRA methods.

Method	Approaches
THERP	Decomposes tasks into subtasks with clear reductionist characteristics (Boring et al., 2005).
ASEP	Proceduralized detail categorizes it as a reductionist method (Boring et al., 2005).
HEART	Decomposes tasks and classifies them, exhibiting clear reductionist features (Williams, 2015).
SLIM-MAUD	Reliance on PSFs categorizes it as reductionist (Boring et al., 2005).
ATHEANA	Hybrid method with holistic tendencies, emphasizing a holistic understanding of context (Boring et al., 2005; De Felice et al., 2012; Dhillon, 2014; Boring & Joe, 2014).
SPAR-H	Assesses context effects atomistically, exhibiting atomistic characteristics (Boring et al., 2005; Dhillon, 2014).
HCR	Explicitly models time-dependent human reliability, considered holistic (Bodberg, 1993; Dhillon, 2014).
CREAM	Basic method exhibits holistic tendencies; extended method shows reductionist tendencies (He, 2008).
INTENT	Labeled as a holistic approach, limiting detailed comparison of communication paths (Zander et al., 1999).

The epistemological diversity among HRA methods is encapsulated in **Table 3**, shedding light on their reductionist or holistic nature. As highlighted by Pencea et al. (2014), challenges arise in bottom-up approaches, leading to the reduction of factors for quantification, potentially decreasing the completeness of contextual depictions. Conversely, top-down approaches face criticism for potential data misinterpretation due to the lack of underlying theories.

This table provides a nuanced overview of the epistemological approaches in various HRA methods, contributing to a comprehensive understanding of their reductionist or holistic orientations.

3. Paradigm Shift in Human Reliability Assessment: Embracing Socio-Technical Systems and Systemic Perspectives

Human Reliability Assessment (HRA) models have transitioned from reductionist to systemic paradigms, underpinned by holistic theories. Since the 1990s, numerous HRA authors have introduced holistic and systemic concepts, catalyzing this paradigm shift. Pioneers such as Rasmussen (Le Coze, 2015), Reason, and Perrow (1999) incorporated notions like degree of freedom, self-organization, barriers, latent failures, and the “system accident” concept, signaling a departure from traditional reductionist views.

In contemporary industrial settings, characterized by collaborative operations and shared responsibilities, a systemic perspective is deemed essential (French, Bedford, Pollard, & Soane, 2011). This contrasts starkly with reductionist approaches that compartmentalize human and machine interactions. System theory-based HRA models strive to capture the intricate causality and complexity of modern socio-technical systems from a broad systemic view (Abraha &

Liyanage, 2012).

- The sociotechnical system concept reflects an integrative vision involving elements with human, technical, and economic features (Tonț, Vlădăreanu, Munteanu, & Tonț, 2009). Emphasizing interdependencies and links, this systemic approach provides a global understanding of economic, social, environmental, and technical aspects. It enables the identification of qualitative and quantitative relationships, ensuring an optimal balance between system components. Integrating both reductionism and holism, this approach becomes instrumental in comprehending complex problems (Ham, Park, & Jung, 2012).

The pivotal concept of “complexity” (Le Coze, 2006; Silberstein & McGeever, 1999; Gertman, 2012) becomes paramount in addressing the behavior of modern socio-technical systems. These systems exhibit emergent properties, self-organization, multiple agents, and adaptive qualities. The dynamism of complexity extends across hardware and software infrastructures, interdependencies, human-machine interactions, organizational behavior, and system adaptability.

Despite the recognition of complexity as a Performance Shaping Factor (PSF) in many HRA models, merely treating it as such does not capture the nuanced interactions, adaptations, coordination, and synchronization inherent in complex systems.

- Cognitive Systems Engineering, pioneered by Hollnagel in the 1980s, represents a significant paradigm shift in human reliability studies (Hollnagel & Woods, 1983). Models like CREAM mark this generational change, emphasizing a move from a focus on the causes of failures to understanding the causes of successful outcomes. Hollnagel’s recent proposition of safety science shifting from Safety-I to Safety-II underscores the need for a systemic view, rejecting decomposable and predictable models (Hollnagel, 2018).

Resonance, conceptual construct introduced by Hollnagel (2017), offer a deeper understanding of accidents. Resonance, derived from physics, highlights the variability in normal performance due to approximate adjustments, revealing the interconnectedness of system functions. Emergence (Goldstein, 1999), a crucial concept in system approaches, aligns with resilience (Hollnagel, Woods, & Leveson, 2006; Woods, 2015) and underscores the non-decomposable nature of socio-technical systems.

- HRA Sociotechnical System-Based Models: Evident in recent research trends is a shift toward systems theory, particularly in the form of systems dynamics modeling. Models like SoTeRiA (Mohaghegh & Mosleh, 2009), CHMS (Boy, 2011), and FRAM (Hollnagel, 2017) are examples of this systemic approach. They illuminate interrelationships and interdependencies among system components, providing a holistic view of complex operations.

These new models, represented in **Table 4**, not only signify this shift but also adopt conceptual constructs from various disciplines. Their graphical representations add descriptive power, simplicity, and visual clarity, reflecting emergent properties within the model.

Table 4. Systemic evidences in HRA models.

Model	Systemic approach evidences	
	Conceptual construct	Interrelationships and interdependencies highlighted
SoTeRiA	General principles for organizational safety causal frameworks. Safety culture. Safety climate.	<p>Definition, elements, antecedents, and outcomes of individual safety performance, safety culture, safety climate, and organizational safety structure and practices are analyzed. It elucidates and schematically represents their interrelations, influences paths, and links.</p> <p>The theoretical relation between organizational culture, structure/practices, and climate is explored. Paths of influence from the organization as a whole, to groups, to individuals, and then from the individual level back up to the organization level safety outcome are highlighted. Links between ‘individual-level PSFs’ and ‘human action’ are established. Organizational practices influencing resources, individuals’ performances, and procedures (‘organizational safety practices’) are examined.</p>
CHMS	Socio-cognitive stability.	<p>Relations between local stability (workload, situation awareness, ability to make appropriate decisions, correct action execution; supported by appropriate redundancies and various kinds of cognitive support) and global stability (appropriateness of functions allocated to agents, pace of information flows, and related coordination; supported by safety net). Socio-cognitive stability is related to socio-cognitive system’s resilience. Safety is reached when human and software agents (as components of human-machine systems) interact to maintain both local and global socio-cognitive stability. Technology and human practice are incrementally co-designed. Levels of difficulty (assessed by time pressure criticality, complexity, and flexibility) are defined to stabilize a multi-agent system in a continuum between passive and active socio-cognitive stability.</p>
FRAM	Resonance	<p>Dependencies and relations among “essential system functions.” Each system function is represented as a hexagon with vertices indicating input, preconditions, resources, time, control, and output, connecting with other essential functions and context. Connections reveal performance variability where resonance may create accident conditions. The systemic view emphasizes how functions depend on each other, and unexpected couplings may suddenly appear. Differences between systemic and holistic ontological approaches are exposed, stating that the description of the overall structure of the system is derived from how connections between functions are specified.</p>
STS	Brownian movement	<p>Hierarchical interactions between different system layers (government, regulators, company, management, staff, and work) and openness to external constraints.</p>
HRO	Mindfulness	<p>Focuses on individual interactions that create collective cognitive behaviors or collective mindfulness. Describes the five collective cognitive processes as the base for achieving mindfulness and subsequently high reliability and performance. Individuals working together with cognitive stability and capabilities to action, anticipate, and adjust their action patterns to understand, collect evidence, detect, evaluate, revise, etc., new or unexpected events.</p>

In **Table 4**, systemic evidence within HRA models is presented, emphasizing the interplay of conceptual constructs and the highlighted interdependencies among system components, showcasing the systemic nature of these models, a perspective crucial in understanding accidents and incidents in complex operations.

4. From Theory to Toolbox: Epistemic Strategies for Practical HRA Advancements

The evolution of HRA models has undergone significant paradigm shifts, reflecting epistemological changes. [Belmonte et al. \(2011\)](#) proposed taxonomies categorizing HRA models into distinct phases: the machine-centered, human-centered, and human-machine system approaches, aligning with the first, second, and third generations of HRA.

The initial machine-centered perspective prioritized technical aspects over human factors. Subsequently, the human-centered approach recognized the indispensable role of human factors in safety. The evolution culminated in the human-machine system approach, integrating both elements, acknowledging the intricate interplay within socio-technical systems. Notably, this systemic perspective becomes particularly relevant when considering Critical Infrastructures (IC) as socio-technical systems, a burgeoning field marked by explicit systemic approaches ([Hollnagel, 2008](#); [Zio, 2016](#)).

Examining three significant tendencies—referred to as the “Ellulian,” “Khunian,” and “Ashbyan” perspectives—provides unique insights into the complex relationship between technology, cognition, and socio-technical systems ([Le Coze, 2015](#)).

The “Ellulian” tendency, grounded in technological determinism, posits that accidents result from technology escaping human control, emphasizing the autonomy of technology and its potential to introduce unforeseen challenges. The “Khunian” tendency adopts constructivist approaches, highlighting the role of cultural-cognitive phenomena in shaping safety outcomes. Lastly, the “Ashbyan” tendency underscores socio-technical systems’ complexity, emergence, and self-organization, framing accidents within a broader systemic context ([De Winter & Dodou, 2014](#)).

A fundamental realization within this paradigm shift is the redefinition of “human error.” No longer viewed merely as a psychological category of human deficiencies, human error is now considered a symptom of systemic vulnerabilities embedded within the organization ([Johannesen, Sarter, Cook, Dekker, Woods, 2012](#)). This conceptual shift moves away from individual-focused explanations, offering a systemic perspective where accidents result from interactions among components violating safety constraints on system design and operation.

Building on the systemic approach, this subsection explores the implications for safety design. While artifacts and systems are traditionally designed through

divided production processes and specialized teams (termed the “positivist approach of design and development” (Boy, 2011)), they function as integrated wholes.

This holistic perspective acknowledges the complex interconnections and emergent properties inherent in socio-technical systems. Safety design, therefore, requires a comprehensive understanding of how different components interact and influence each other. By considering the system as a whole, safety design can address latent vulnerabilities and enhance overall system resilience (Pariès, 2012).

Resilience emerges as a pivotal conceptual construct, extensively researched within safety science (Hollnagel, Woods, & Leveson, 2006; Woods, 2015) and other disciplines (Bhamra, Dani, & Burnard, 2011). Mohaghegh and Mosleh (2009) describe three emergent processes within socio-technical systems: “homogeneity,” “social interaction,” and “leadership.” Additionally, Rasmussen’s “defense in depth fallacy” (Le Coze, 2013) accentuates the interdependencies among actors and the repercussions of local defenses violations. This systemic perspective posits accidents as emergent phenomena, a consequence of intricate and nonlinear interactions among system components (Abraha & Liyanage, 2012). Incorporating these insights into the discussion enriches our understanding of resilience within the context of systemic approaches in Human Reliability Analysis.

The evolving paradigms in HRA models have profound implications for conceptualizing accidents and understanding the interconnectedness of elements within socio-technical systems. This shift paves the way for advanced Human Reliability Assessment models grounded in a systemic understanding of the intricate relationships at play. Developing these advanced models requires interdisciplinary collaboration, integrating insights from psychology, engineering, and organizational studies. By embracing a holistic perspective, future HRA models can more effectively capture the complexity of modern socio-technical systems, contributing to enhanced safety and risk management practices.

The emergence of HRA originated from the engineering need to measure and predict human errors in critical industrial plants. As such, HRA studies inherently adopted pragmatism and empiricism associated with the efficiency and measurement characteristic of engineering studies (Melles, 2008). Contrarily, HRA can also be considered an applied field within disciplines concerned with human phenomena, such as anthropology, psychology, or sociology. In contrast to these disciplines, HRA lacks a single theoretical perspective, exhibiting a variety of philosophical assumptions guiding model and method development. The epistemological analysis presented in this paper aims to uncover these hidden assumptions, promoting greater awareness and debate within the HRA community.

Philosophical assumptions and theoretical frameworks, also known as “metatheory” (Hillix & L’Abate, 2012), play a crucial role in guiding how HRA mod-

els and theories are constructed. However, these assumptions often remain concealed from view, with researchers possibly unaware of their metatheoretical stances. The relationship between epistemology and method is rarely articulated, as publications tend to emphasize methods rather than the entire construction of the research process (Darlaston-Jones, 2007; Hillix & L'Abate, 2012). The epistemological analysis in this paper seeks to discuss the range of philosophical assumptions or metatheories that underpin different HRA models, encouraging greater debate and mindfulness around these assumptions and their practical consequences (Abraha & Liyanage, 2012). This move aligns with the broader trend of practice fields seeking to legitimize their academic status (Melles, 2008).

Epistemological analysis serves four key functions for research and practice in HRA (Winsberg, 1999). First, it facilitates usability analysis, making underlying decisions explicit for knowledge reuse in future studies. Second, it acts as a research aid, enabling researchers to plan, address, and structure future research aligned with their perspectives. Third, it guides practice by assisting users in understanding the assumptions underlying methods and making informed choices when designing interventions. Lastly, it facilitates communication within and between development teams by providing a common language and vocabulary for discourse.

The epistemological analysis presented in this paper constructs a generic typology representing the main philosophical assumptions underpinning different HRA models (Mingers, 2003). This awareness is essential for researchers to understand and improve their models, argue against epistemologically based criticism, and align their research perspectives with their research traditions and philosophical assumptions (Mitcham, 1998).

Theoretical foundations presented in this paper reflect diverse positions with regard to ontological, gnosiological, and methodological aspects. Choosing one over the other can be linked to personal convictions, beliefs, and various institutional contexts (Orlikowski & Baroudi, 1991; Le Coze, 2022). Despite the increasing systemic tendency in recent HRA methods, there are no definite answers, as debates around reductionist, holistic, and systemic approaches persist in philosophy. HRA researchers are encouraged to follow these debates to incorporate the latest theoretical advances into their work.

The epistemological assessment carried out in this study holds profound practical implications for the development and enhancement of HRA models. By rigorously evaluating the philosophical assumptions, this assessment contributes to a deeper understanding of how these assumptions impact the reliability and effectiveness of HRA in real-world applications. It encourages researchers to engage in critical discourse, improve model robustness, and foster innovation within the field. Epistemological awareness is paramount for ensuring the relevance, effectiveness, and evolution of HRA in the face of dynamic challenges within socio-technical systems.

Certainly! Here's a suggestion for providing more detailed suggestions for future establishments of the new HRA model:

Drawing insights from the comprehensive examination of existing HRA models and their underlying epistemological assumptions, this study lays the groundwork for future advancements in the field. The analysis prompts specific recommendations for the development of a new HRA model that transcends current limitations. Firstly, emphasizing explicit articulation of philosophical assumptions should become a foundational practice in model construction. Integrating diverse epistemological perspectives, such as reductionism, holism, and systemism, into the fabric of the model will foster a more comprehensive understanding of human reliability phenomena. Secondly, fostering interdisciplinary collaboration is paramount. In future HRA model development, incorporating perspectives from psychology, engineering, and organizational studies can enrich the model's robustness and applicability. Thirdly, the study encourages a paradigm shift toward a more balanced integration of theoretical considerations, ensuring that HRA models not only serve practical needs but also contribute to the broader philosophical discourse within the discipline. By considering these detailed suggestions, future HRA models can aspire to be more nuanced, adaptable, and reflective of the evolving epistemological landscape in human reliability studies.

5. Conclusion and Discussion

The comprehensive assessment of epistemological foundations in Human Reliability Analysis (HRA) models has revealed nuanced insights into the evolution of paradigms and philosophical assumptions within the discipline. The reductionist paradigm, rooted in historical perspectives, was critically examined for its implications on understanding human error and influencing safety design. The shift towards holistic theories underscored the importance of integrating human, technical, and organizational facets within modern socio-technical systems. Furthermore, the exploration of systemic approaches marked a paradigmatic evolution towards a broader, interdependent view of accidents and reliability. The integrative summary of assessments across these paradigms contributes to a multifaceted understanding of HRA models' philosophical underpinnings. Recognizing the strengths and limitations within each paradigm informs the ongoing discourse on enhancing the discipline's scientific validity and practical application. As HRA strives to establish itself as a scientific discipline, this study advocates for continued assessment and scrutiny, fostering a paradigm shift that rethinks the epistemological bases of models. Such a shift is imperative for advancing HRA, ensuring its relevance in comprehending, predicting, and improving human performance and system safety in complex operational settings.

As we traverse the evolving landscape of Human Reliability Analysis (HRA), this paper has ventured beyond the practical realm to scrutinize its epistemological underpinnings. In doing so, it has transcended the conventional narrative, emphasizing the importance of conscious reflection on philosophical assump-

tions in HRA model and method construction.

The original contribution of this paper lies in its meticulous examination of influential HRA models through the lenses of reductionism, holism, and systemism. While acknowledging the impact of other epistemological approaches, our focus on these three perspectives has shed light on the implicit assumptions that have steered HRA's trajectory.

Despite HRA's continuous efforts to enhance models for predicting human performance and ensuring system safety, recent incidents highlight persistent challenges. This paper contends that epistemological assumptions, whether explicit or implicit, have played a pivotal role in shaping these models. Even post-development epistemological analyses have revealed internal issues that may contribute to limitations in addressing emerging accident scenarios.

Contrary to the traditional approach of building upon existing models, we advocate a paradigm shift: a step back to reevaluate epistemological foundations. The proposition is to embark on the creation of an entirely new HRA model, guided by robust ontological, gnosiological, and methodological considerations. This approach aligns with the core principles of scientific theories—plausibility, explanatory adequacy, interpretability, simplicity, descriptive adequacy, and generalizability.

Engaging in this crucial epistemological discourse holds profound significance in the realm of human reliability studies within engineering sciences. It not only encourages a deeper understanding of the philosophical underpinnings but also sets the stage for innovative advancements. As HRA continues to navigate dynamic challenges in socio-technical systems, a conscious integration of epistemological considerations is paramount for scientific progress and the evolution of the discipline.

Acknowledgements

Authors acknowledge the financial support received from the Secretaría de Investigaciones Internacional y Posgrado (SIIP) of the National University of Cuyo, under project number 06/B029-T1.

Additionally, authors extend sincere gratitude to the Facultad de Ciencias Económicas (Faculty of Economic Sciences) at the National University of Cuyo for their invaluable support throughout the duration of this project. The academic environment and resources provided by the institution played a crucial role in facilitating the research process and enhancing the quality of the outcomes. This acknowledgment reflects the commitment of the faculty to fostering academic excellence and contributing to the advancement of knowledge.

We also express our sincere appreciation to the Secretaría de Investigación (Research Secretariat) of ESEADE (Escuela Superior de Economía y Administración de Empresas) for their invaluable support throughout the duration of this project. Their commitment to fostering research and academic endeavors has been instrumental in the successful completion of our work.

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

- Abraha, H. H., & Liyanage, J. P. (2012). Review of Theoretical Foundations for Risk Minimal Operations in Complex Sociotechnical Systems: The Role of Human Error. In W. Lee, B. Choi, L. Ma, & J. Mathew (Eds.), *Proceedings of the 7th World Congress on Engineering Asset Management (WCEAM 2012)* (pp. 1-16). Springer International Publishing. https://doi.org/10.1007/978-3-319-06966-1_1
- Abraha, H. H., & Liyanage, J. P. (2015). Managing Modern Sociotechnical Systems: New Perspectives on Human-Organization—Technological Integration in Complex and Dynamic Environments. In P. Tse et al. (Eds.), *Engineering Asset Management—Systems, Professional Practices and Certification* (pp. 1109-1123). Springer International Publishing. https://doi.org/10.1007/978-3-319-09507-3_94
- Anderson, P. A. (1972). More Is Different: Broken Symmetry and the Nature of the Hierarchical Structure of Science. *Science*, *177*, 393-396. <https://doi.org/10.1126/science.177.4047.393>
- Ayala, F. J. (1987). Biological Reductionism. In F. E. Yates et al. (Eds.), *Self-Organizing Systems* (pp. 315-324). Springer US. https://doi.org/10.1007/978-1-4613-0883-6_17
- Bell, J., & Holroyd, J. (2009). *Review of Human Reliability Assessment Methods*. Health and Safety Laboratory.
- Belmonte, F., Schön, W., Heurley, L., & Capel, R. (2011). Interdisciplinary Safety Analysis of Complex Socio-Technological Systems Based on the Functional Resonance Accident Model: An Application to Railway Traffic Supervision. *Reliability Engineering & System Safety*, *96*, 237-249. <https://doi.org/10.1016/j.res.2010.09.006>
- Bhamra, R., Dani, S., & Burnard, K. (2011). Resilience: The Concept, a Literature Review and Future Directions. *International Journal of Production Research*, *49*, 5375-5393. <https://doi.org/10.1080/00207543.2011.563826>
- Block, N., & Stalnaker, R. (1999). Conceptual Analysis, Dualism, and the Explanatory Gap. *The Philosophical Review*, *108*, 1-46. <https://doi.org/10.2307/2998259>
- Bodsberg, L. (1993). Comparative Study of Quantitative Models for Hardware, Software and Human Reliability Assessment. *Quality and Reliability Engineering International*, *9*, 501-518. <https://doi.org/10.1002/qre.4680090607>
- Boring, R. L., & Joe, J. C. (2014). *Task Decomposition in Human Reliability Analysis* (No. InL/CON-14-31656), Idaho National Lab. (INL).
- Boring, R. L., Gertman, D. I., Joe, J. C., & Marble, J. L. (2005). Human Reliability Analysis in the US Nuclear Power Industry: A Comparison of Atomistic and Holistic Methods. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *49*, 1815-1819. <https://doi.org/10.1177/154193120504901913>
- Boy, G. A. (2011). Design for Safety: A Cognitive Engineering Approach. In Y. Barnard (Ed.), *Safety of Intelligent Driver Support Systems. Design, Evaluation and Social Perspectives*. CRC Press.
- Brewer, J. D. (2006). Multidimensional Confusability Matrices Enhance Systematic Analysis of Unsafe Actions and Human Failure Events Considered in PSAs of Nuclear Power Plants (PSAM-0334). In M. G. Stamatelatos, & H. S. Blackman (Eds.), *Proceedings of the Eighth International Conference on Probabilistic Safety Assessment & Management (PSAM)*. ASME Press.

- Bunge, M. (2000). Ten Modes of Individualism—None of Which Works—And Their Alternatives. *Philosophy of the Social Sciences*, 30, 384-406.
<https://doi.org/10.1177/004839310003000303>
- Bunge, M. (2003). *Emergence and Convergence: Qualitative Novelty and the Unity of Knowledge*. University of Toronto Press. <https://doi.org/10.3138/9781442674356>
- Cacciabue, P. (2000). Human Factors Impact on Risk Analysis of Complex Systems. *Journal of Hazardous Materials*, 71, 101-116.
[https://doi.org/10.1016/S0304-3894\(99\)00074-6](https://doi.org/10.1016/S0304-3894(99)00074-6)
- Darlaston-Jones, D. (2007). Making Connections: The Relationship between Epistemology and Research Methods. *Australian Community Psychologist*, 19, 19-27.
- De Felice, F., Petrillo, A., Carlomusto, A., & Ramondo, A. (2012). Human Reliability Analysis: A Review of the State of the Art. *IRACST—International Journal of Research in Management & Technology*, 2, 35-41.
- De Winter, J. C., & Dodou, D. (2014). Why the Fitts List Has Persisted Throughout the History of Function Allocation. *Cognition, Technology & Work*, 16, 1-11.
<https://doi.org/10.1007/s10111-011-0188-1>
- Dekker, S. W. (2005). *Ten Questions about Human Error: A New View of Human Factors and System Safety*. Lawrence Erlbaum Associates. <https://doi.org/10.1201/b12474>
- Dekker, S., & Hollnagel, E. (2004). Human Factors and Folk Models. *Cognition, Technology & Work*, 6, 79-86. <https://doi.org/10.1007/s10111-003-0136-9>
- Dhillon, B. S. (2014). Specific Human Reliability Analysis Methods for Nuclear Power Plants. In *Human Reliability, Error, and Human Factors in Power Generation* (pp. 65-79). Springer International Publishing.
https://doi.org/10.1007/978-3-319-04019-6_5
- Di Pasquale, V., Iannone, R., Miranda, S., & Riemma, S. (2013). An Overview of Human Reliability Analysis Techniques in Manufacturing Operations. *Operations Management*, 9, 978-953. <https://doi.org/10.5772/55065>
- Farrington-Darby, T., & Wilson, J. R. (2006). The Nature of Expertise: A Review. *Applied Ergonomics*, 37, 17-32. <https://doi.org/10.1016/j.apergo.2005.09.001>
- Forester, J., Bley, D., Cooper, S., Lois, E., Siu, N., Kolaczowski, A., & Wreathall, J. (2004). Expert Elicitation Approach for Performing ATHEANA Quantification. *Reliability Engineering & System Safety*, 83, 207-220. <https://doi.org/10.1016/j.res.2003.09.011>
- French, S., Bedford, T., Pollard, S. J., & Soane, E. (2011). Human Reliability Analysis: A Critique and Review for Managers. *Safety Science*, 49, 753-763.
<https://doi.org/10.1016/j.ssci.2011.02.008>
- Gertman, D. I. (2012). Complexity: Application to Human Performance Modeling and HRA for Dynamic Environments. In *5th International Symposium on Resilient Control Systems (ISRCs)* (pp. 19-24). IEEE. <https://doi.org/10.1109/ISRCs.2012.6309287>
- Goldstein, J. (1999). Emergence as a Construct: History and Issues. *Emergence*, 1, 49-72.
https://doi.org/10.1207/s15327000em0101_4
- Gregoriades, A., & Sutcliffe, A. (2008). Workload Prediction for Improved Design and Reliability of Complex Systems. *Reliability Engineering and System Safety*, 93, 530-549.
<https://doi.org/10.1016/j.res.2007.02.001>
- Ham, D.-H., Park, J., & Jung, W. (2012). Model-Based Identification and Use of Task Complexity Factors of Human Integrated Systems. *Reliability Engineering & System Safety*, 100, 33-47. <https://doi.org/10.1016/j.res.2011.12.019>
- He, X. W. (2008). A Simplified CREAM Prospective Quantification Process and Its Ap-

- plication. *Reliability Engineering & System Safety*, 93, 298-306.
<https://doi.org/10.1016/j.res.2006.10.026>
- Hillix, W. A., & L'Abate, L. (2012). The Role of Paradigms in Science and Theory Construction. In L. L'Abate (Ed.), *Paradigms in Theory Construction* (pp. 3-17). Springer.
https://doi.org/10.1007/978-1-4614-0914-4_1
- Hollnagel, E. (2008). Critical Information Infrastructures: Should Models Represent Structures or Functions? In M. D. Harrison, & M A. Sujan (Eds.), *Computer Safety, Reliability, and Security. SAFECOMP 2008. Lecture Notes in Computer Science* (pp. 1-4). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-540-87698-4_1
- Hollnagel, E. (2017). *FRAM: The Functional Resonance Analysis Method: Modelling Complex Socio-Technical Systems*. CRC Press.
- Hollnagel, E. (2018). *Safety-I and Safety-II: The Past and Future of Safety Management*. CRC Press. <https://doi.org/10.1201/9781315607511>
- Hollnagel, E., & Woods, D. D. (1983). Cognitive Systems Engineering: New Wine in New Bottles. *International Journal of Man-Machine Studies*, 18, 583-600.
[https://doi.org/10.1016/S0020-7373\(83\)80034-0](https://doi.org/10.1016/S0020-7373(83)80034-0)
- Hollnagel, E., Woods, D. D., & Leveson, N. (2006). *Resilience Engineering: Concepts and Precepts*. Ashgate Publishing, Ltd.
- Johannesen, L., Sarter, N., Cook, R., Dekker, S., & Woods, D. D. (2012). *Behind Human Error*. Ashgate Publishing, Ltd.
- Kim, I. S. (2001). Human Reliability Analysis in the Man-Machine Interface Design Review. *Annals of Nuclear Energy*, 28, 1069-1081.
[https://doi.org/10.1016/S0306-4549\(00\)00120-1](https://doi.org/10.1016/S0306-4549(00)00120-1)
- Kirwan, B. (2017). *A Guide to Practical Human Reliability Assessment*. CRC Press.
<https://doi.org/10.1201/9781315136349>
- Latorella, K. A., & Prabhu, P. V. (2000). A Review of Human Error in Aviation Maintenance and Inspection. *International Journal of Industrial Ergonomics*, 26, 133-161.
[https://doi.org/10.1016/S0169-8141\(99\)00063-3](https://doi.org/10.1016/S0169-8141(99)00063-3)
- Le Coze, J. C. (2006). Safety and Security in the Light of Complexity. Uncertainty & Qualification of Systems Analysis. In *3 International Symposium on Systems & Human Science "Complex Systems Approaches for Safety Security and Reliability"* (p. 12). European Communities.
- Le Coze, J. C. (2013). New Models for New Times. An Anti-Dualist Move. *Safety Science*, 59, 200-218. <https://doi.org/10.1016/j.ssci.2013.05.010>
- Le Coze, J. C. (2015). Reflecting on Jens Rasmussen's Legacy. A Strong Program for a Hard Problem. *Safety Science*, 71, 123-141. <https://doi.org/10.1016/j.ssci.2014.03.015>
- Le Coze, J. C. (2022). The 'New View' of Human Error. Origins, Ambiguities, Successes and Critiques. *Safety Science*, 154, Article 105853.
<https://doi.org/10.1016/j.ssci.2022.105853>
- Machado Susseret, N. (2005). *Fiabilidad humana en los sistemas de salud y seguridad laboral de las organizaciones*. Centro de Documentación, Facultad de Ciencias Económicas y Sociales, Universidad Nacional de Mar del Plata.
- Melles, G. (2008). An Enlarged Pragmatist Inquiry Paradigm for Methodological Pluralism in Academic Design Research. *Artifact*, 2, 3-11.
<https://doi.org/10.1080/17493460802276786>
- Mingers, J. (2003). A Classification of the Philosophical Assumptions of Management Science Methods. *The Journal of the Operational Research Society*, 54, 559-570.

- <https://doi.org/10.1057/palgrave.jors.2601436>
- Mitcham, C. (1998). The Importance of Philosophy to Engineering. *Teorema: Revista Internacional de Filosofía*, XVII, 27-47.
- Mohaghegh, Z., & Mosleh, A. (2009). Incorporating Organizational Factors into Probabilistic Risk Assessment of Complex Socio-Technical Systems: Principles and Theoretical Foundations. *Safety Science*, 47, 1139-1158.
<https://doi.org/10.1016/j.ssci.2008.12.008>
- Mosleh, A. H., & Chang, Y. (2004). Model-Based Human Reliability Analysis: Prospects and Requirements. *Reliability Engineering and System Safety*, 83, 241-253.
<https://doi.org/10.1016/j.res.2003.09.014>
- Mosleh, A., & Chang, Y. H. (2007). Cognitive Modelling and Dynamic Probabilistic Simulation of Operating Crew Response to Complex System Accidents—Part 1. *Reliability Engineering and System Safety*, 92, 997-1013.
<https://doi.org/10.1016/j.res.2006.05.014>
- Munger, S., Smith, R., & Payne, D. (1962). *An Index of Electronic Equipment Operability, Data Store*. AIR-C43-1/62-RP, America Institute for Research.
<https://doi.org/10.21236/AD0607161>
- Orlikowski, W. J., & Baroudi, J. J. (1991). Studying Information Technology in Organizations: Research Approaches and Assumptions. *Information Systems Research*, 2, 1-28.
<https://doi.org/10.1287/isre.2.1.1>
- Pariès, J. (2012) A Complexity-Of-Strategic-Behavior Comparison between Schulze's Rule and Ranked Pairs. *26th AAAI Conference on Artificial Intelligence*, 26, 1429-1435.
<https://doi.org/10.1609/aaai.v26i1.8258>
- Pencea, J. M., Ostroff, C., Keec, E., Yilmazd, F., Grantome, R., & Johnsonf, D. (2014). Toward Monitoring Organizational Safety Indicators by Integrating Probabilistic Risk Assessment, Socio-Technical Systems Theory, and Big Data Analytics. In *Proceedings of 12th International Topical Meeting on Probabilistic Safety Assessment and Analysis* (549-564).
- Perrow, C. (1999). *Normal Accidents: Living with High-Risk Technologies*. Princeton University Press. <https://doi.org/10.1515/9781400828494>
- Pyy, P. (2000). *Human Reliability Analysis Methods for Probabilistic Safety Assessment (Thesis)*. VTT PUBLICATIONS.
- Richei, A., Hauptmanns, U., & Unger, H. (2001). The Human Error Rate Assessment and Optimizing System HEROS—A New Procedure for Evaluating and Optimizing the Man-Machine Interface in PSA. *Reliability Engineering and System Safety*, 72, 153-164.
[https://doi.org/10.1016/S0951-8320\(01\)00005-9](https://doi.org/10.1016/S0951-8320(01)00005-9)
- Silberstein, M., & McGeever, J. (1999). The Search for Ontological Emergence. *The Philosophical Quarterly*, 49, 201-214. <https://doi.org/10.1111/1467-9213.00136>
- Sornette, D., Maillart, T., & Kröger, W. (2013). Exploring the Limits of Safety Analysis in Complex Technological Systems. *International Journal of Disaster Risk Reduction*, 6, 59-66. <https://doi.org/10.1016/j.ijdrr.2013.04.002>
- Stanton, N. A. (2016). Distributed Situation Awareness. *Theoretical Issues in Ergonomics Science*, 17, 1-7. <https://doi.org/10.1080/1463922X.2015.1106615>
- Tonț, G., Vlădăreanu, L., Munteanu, R. A., & Tonț, D. G. (2009). Some Aspects Regarding Human Error Assessment in Resilient Socio-Technical Systems. In *WSEAS International Conference on Mathematical and Computational Methods in Science and Engineering* (pp. 139-144). World Scientific and Engineering Academy and Society.

- Udehn, L. (2002). The Changing Face of Methodological Individualism. *Annual Review of Sociology*, 28, 479-507. <https://doi.org/10.1146/annurev.soc.28.110601.140938>
- White, D. (1995). Application of Systems Thinking to Risk Management: A Review of the Literature. *Management Decision*, 33, 35-45. <https://doi.org/10.1108/EUM0000000003918>
- Williams, J. C. (2015). Heart—A Proposed Method for Achieving High Reliability in Process Operation by Means of Human Factors Engineering Technology. *Safety and Reliability*, 35, 5-25.
- Winsberg, E. (1999). Sanctioning Models: The Epistemology of Simulation. *Science in Context*, 12, 275-292. <https://doi.org/10.1017/S0269889700003422>
- Woods, D. (2015). Four Concepts for Resilience and the Implications for the Future of Resilience Engineering. *Reliability Engineering and System Safety*, 141, 5-9. <https://doi.org/10.1016/j.res.2015.03.018>
- Zander, R. M., Pamme, H., & Laasko, L. R. (1999). *Reliability Data Assessment*. OCDE.
- Zio, E. (2009). Reliability Engineering: Old Problems and New Challenges. *Reliability Engineering and System Safety*, 94, 125-141. <https://doi.org/10.1016/j.res.2008.06.002>
- Zio, E. (2016). Challenges in the Vulnerability and Risk Analysis of Critical Infrastructures. *Reliability Engineering and System Safety*, 152, 137-150. <https://doi.org/10.1016/j.res.2016.02.009>