UniSC Australia

Centre for Human Factors and Sociotechnical Systems

Playing well together: Sociotechnical system design in the age of AI

Associate Professor Gemma Read

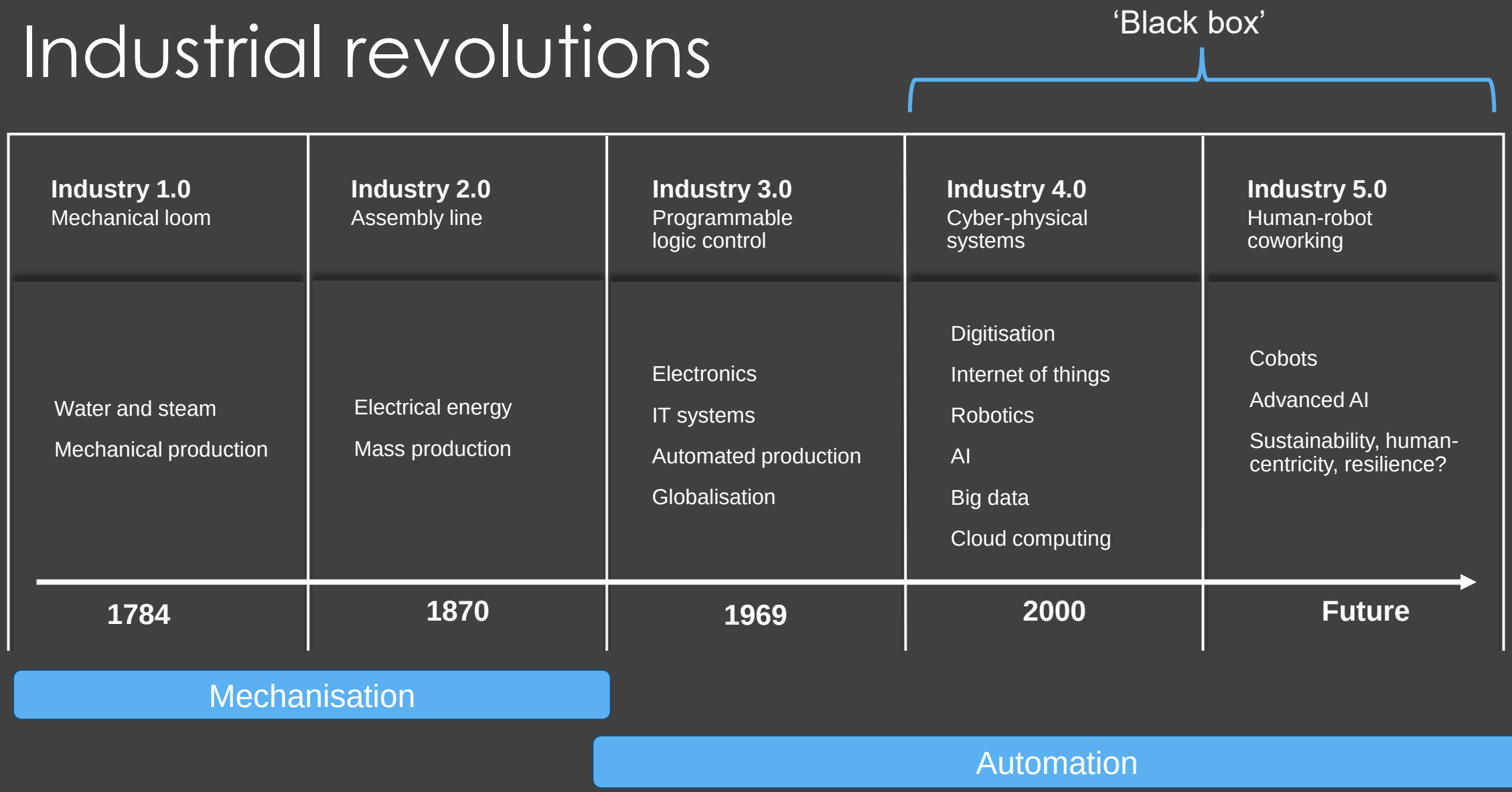
Centre for Human Factors and Sociotechnical Systems

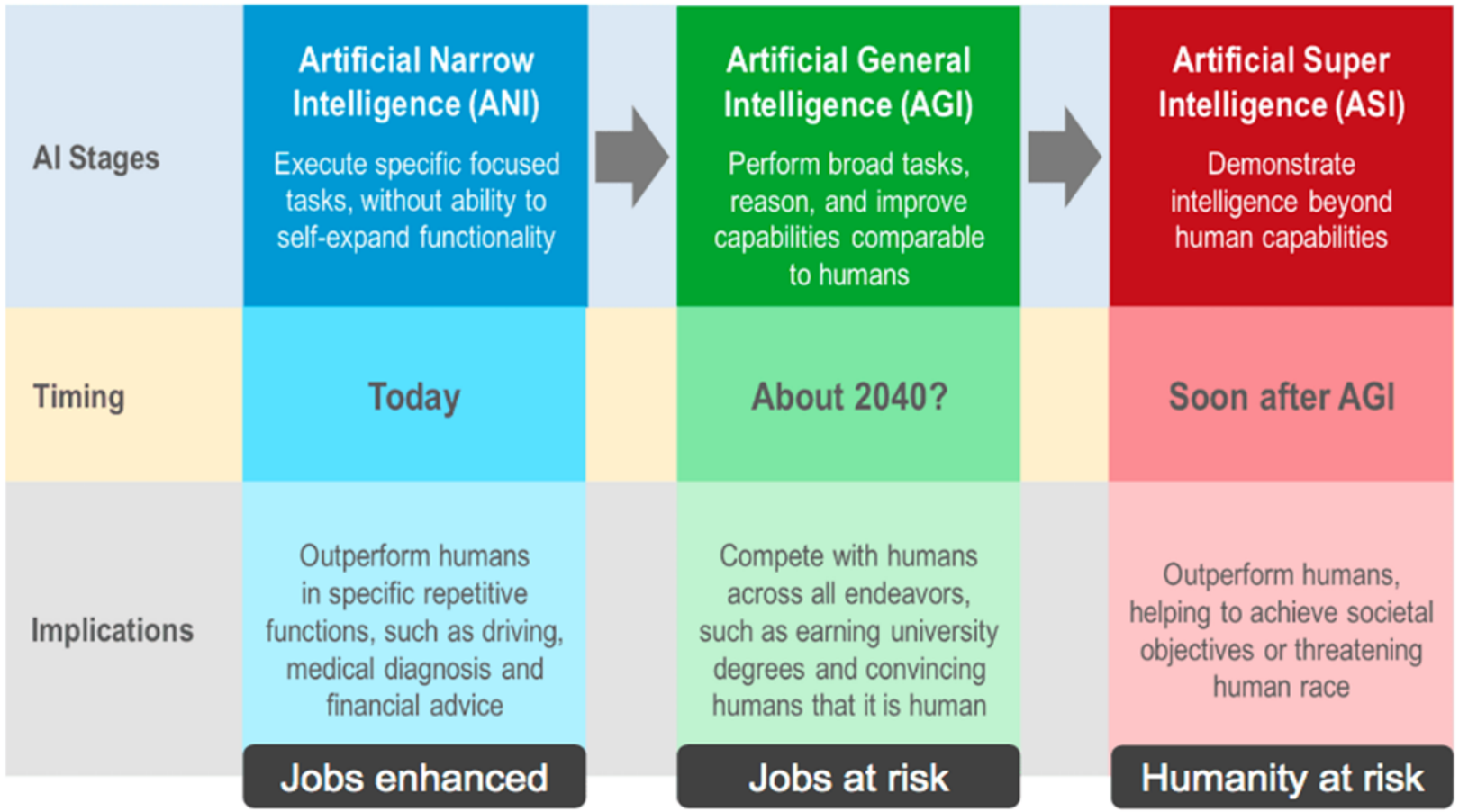
University of the Sunshine Coast

gread@usc.edu.au

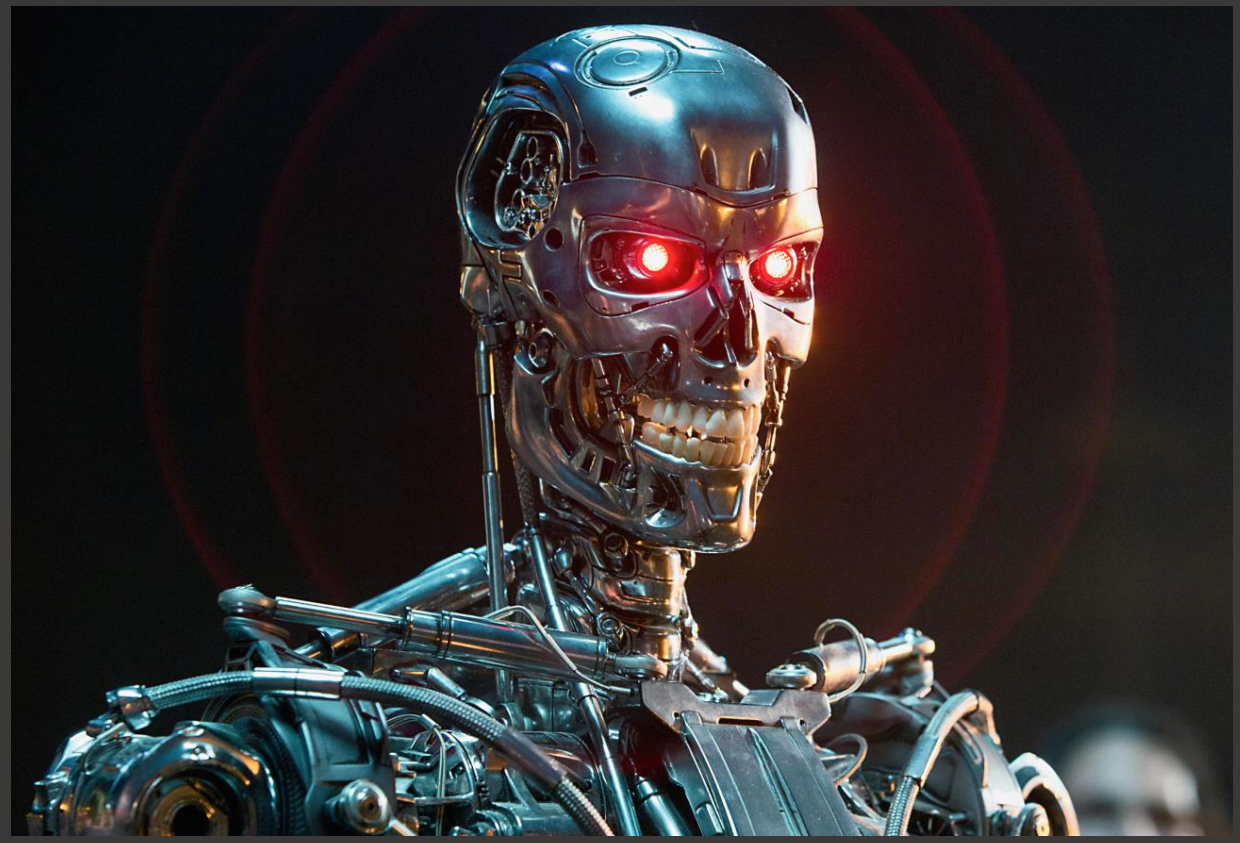
Credit: Adobe Firefly

Industrial revolutions



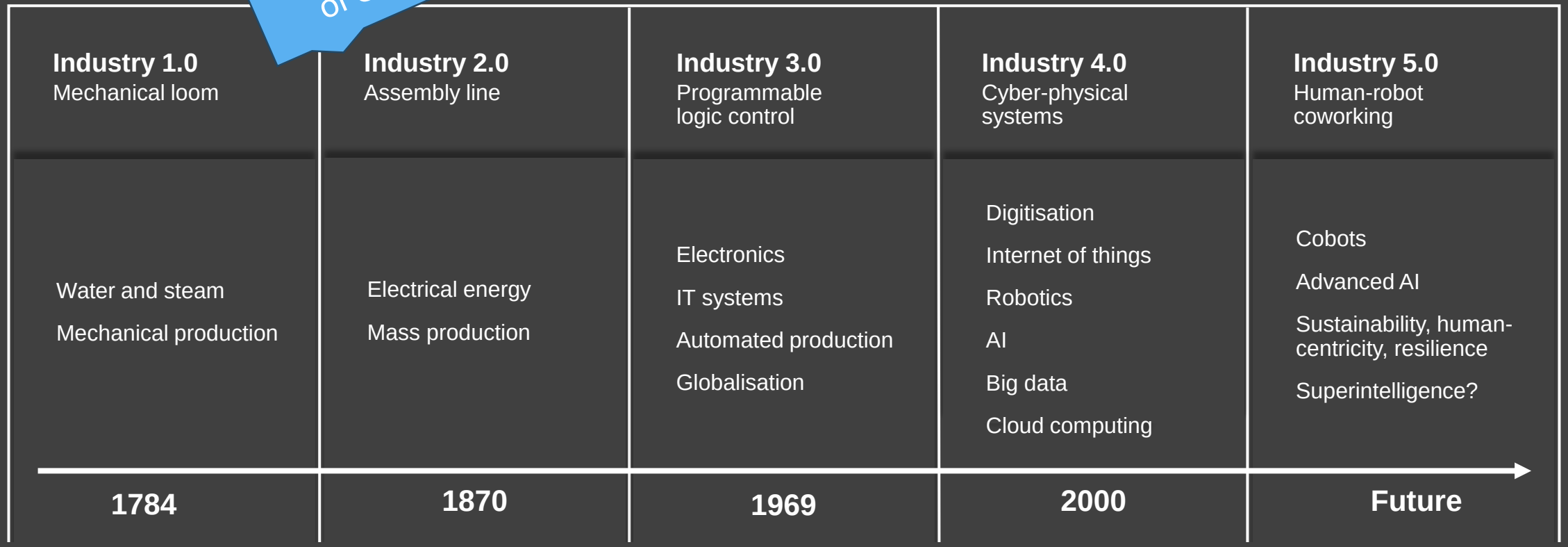


Lawtomated, 2019

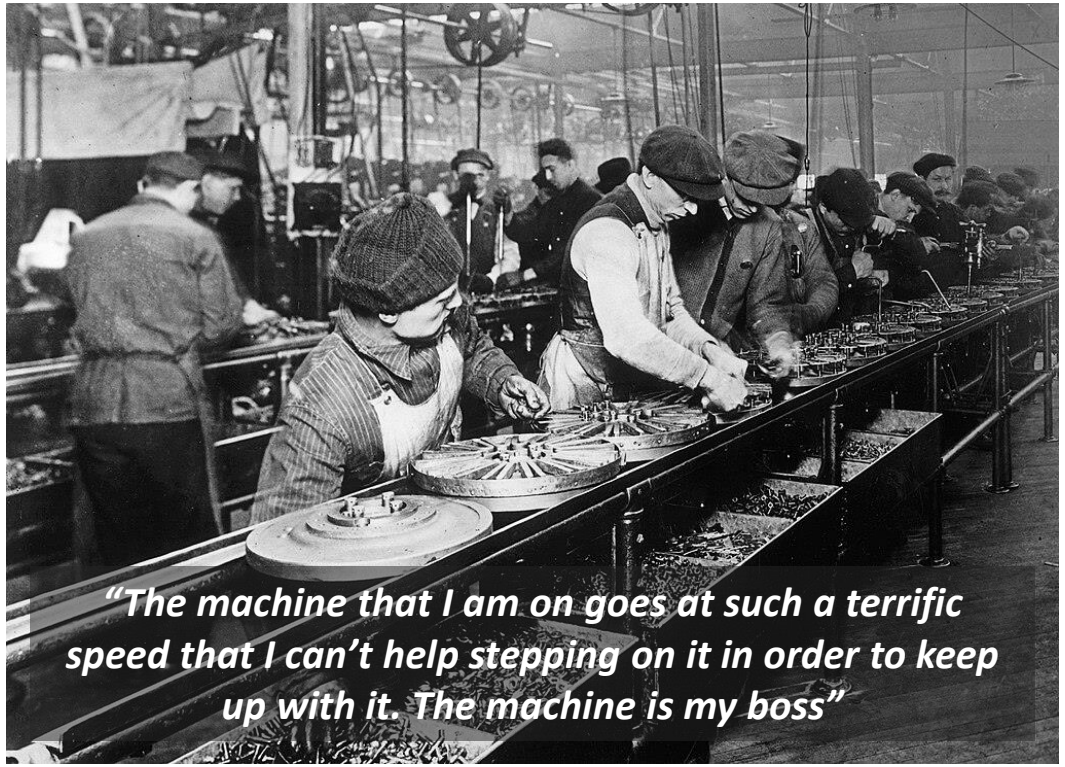


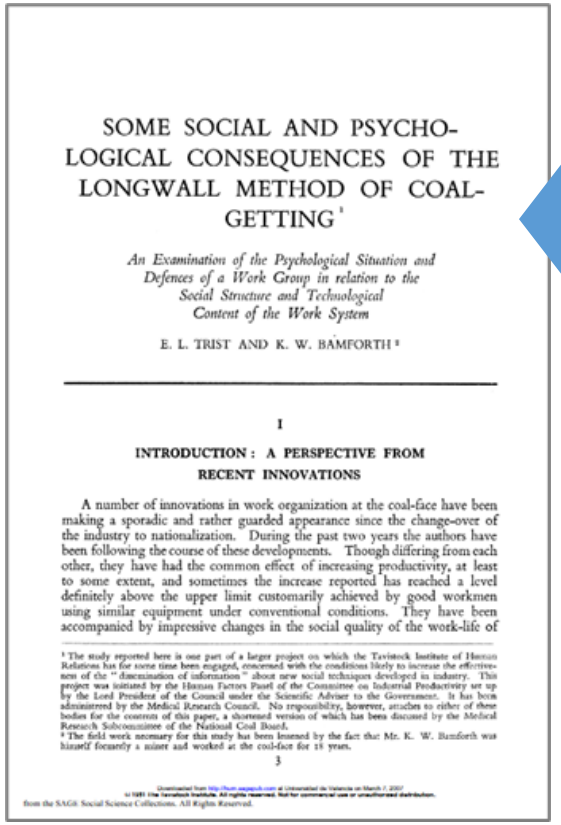
Industrial revolutions

Emergence of STS



Consequences of mechanisation





SOME SOCIAL AND PSYCHO-LOGICAL CONSEQUENCES OF THE LONGWALL METHOD OF COAL-GETTING[1]

An Examination of the Psychological Situation and Defences of a Work Group in relation to the Social Structure and Technological Content of the Work System

E. L. TRIST AND K. W. BAMFORTH[2]

I

INTRODUCTION: A PERSPECTIVE FROM RECENT INNOVATIONS

A number of innovations in work organization at the coal-face have been making a sporadic and rather guarded appearance since the change-over of the industry to nationalization. During the past two years the authors have been following the course of these developments. Though differing from each other, they have had the common effect of increasing productivity, at least to some extent, and sometimes the increase reported has reached a level definitely above the upper limit customarily achieved by good workmen using similar equipment under conventional conditions. They have been accompanied by impressive changes in the social quality of the work-life of

[1] The study reported here is one part of a larger project on which the Tavistock Institute of Human Relations has for some time been engaged, concerned with the conditions likely to increase the effectiveness of the "dissemination of information" about new social techniques developed in industry. This project was initiated by the Human Factors Panel of the Committee on Industrial Productivity set up by the Lord President of the Council under the Scientific Adviser to the Government. It has been administered by the Medical Research Council. No responsibility, however, attaches to either of these bodies for the contents of this paper, a shortened version of which has been discussed by the Medical Research Subcommittee of the National Coal Board.

[2] The field work necessary for this study has been lessened by the fact that Mr. K. W. Bamforth was himself formerly a miner and worked at the coal-face for 18 years.

3

- Fragmentation of work groups
- Loss of personal relationships
- Specialisation & skill degradation
- Increased conflict & retaliation
- Mistrust
- 'Psychosomatic' disorders
- Absenteeism

Rage against the machine: Humanising work systems

Credit: Simon Baddely, Tavistock Researchers, 1969

Sociotechnical systems theory

- Response to increasing technology in workplaces, and Taylorism / scientific management approaches
- Underpinned by:
    - industrial democracy
    - participatory design
    - humanistic values
    - idea of meaningful work
- Organisations as open systems affected by their environment, must respond to a changing context and disturbances
- Adaptive capacity through joint optimisation of technological and social aspects
- Theory evolved a set of values and principles over time

Review of STS studies found improvements relating to productivity (87%), safety (88%), attitudes (94%), absenteeism (81%) and quality (97%) (Pasmore et al., 1982)

Volvo Uddevalla automotive plant

Synthesis of STS values & principles

Ergonomics, 2015
Vol. 58, No. 5, 822–851, http://dx.doi.org/10.1080/00140139.2014.980335

Taylor & Francis

Designing sociotechnical systems with cognitive work analysis: putting theory back into practice

Gemma J.M. Read[a]*, Paul M. Salmon[b], Michael G. Lenné[a] and Neville A. Stanton[c]

[a]*Human Factors Group, Monash Injury Research Institute, Accident Research Centre, Monash University, Clayton, Victoria, Australia;* [b]*University of the Sunshine Coast Accident Research, University of the Sunshine Coast, Maroochydore, Queensland, Australia;* [c]*Transportation Research Group, Civil, Maritime, Environmental Engineering & Science Unit, The University of Southampton, Highfield, Southampton SO17 1BJ, UK*

(*Received 12 March 2014; accepted 19 October 2014*)

Cognitive work analysis (CWA) is a framework of methods for analysing complex sociotechnical systems. However, the translation from the outputs of CWA to design is not straightforward. Sociotechnical systems theory provides values and principles for the design of sociotechnical systems which may offer a theoretically consistent basis for a design approach for use with CWA. This article explores the extent to which CWA and sociotechnical systems theory offer complementary perspectives and presents an abstraction hierarchy (AH), based on a review of literature, that describes an 'optimal' CWA and sociotechnical systems theory design system. The optimal AH is used to assess the extent to which current CWA-based design practices, uncovered through a survey of CWA practitioners, aligns with sociotechnical systems theory. Recommendations for a design approach that would support the integration of CWA and sociotechnical systems theory design values and principles are also derived.

Practitioner Summary: Cognitive work analysis (CWA) is commonly used by ergonomics practitioners for evaluating complex systems and informing the development of design improvements. Despite this, translation from analysis to design is not straightforward. Building upon synergies between CWA and sociotechnical systems design principles, recommendations for a design toolkit are specified.

Keywords: cognitive work analysis; sociotechnical systems theory; system design; complex systems

1. Introduction

Cognitive work analysis (CWA) is a commonly used framework of methods (Salmon et al. 2010) that aims to improve system design (Vicente 1999). While CWA has been used in many design applications (e.g. Bisantz et al. 2003; Naikar et al. 2003; Jenkins, Salmon, et al. 2010; Stanton and McIlroy 2012), like all human factors/ergonomics (HFE) analysis methods, the outputs of CWA provide information to support design activities rather than yielding concrete designs per se. The analysis outputs provide recommendations for various types of interventions, rather than specifying a system fully (Lintern 2005). Furthermore, there has been limited evidence in the open literature of the direct application of CWA outputs in design (Salmon et al. 2010), and the majority of those available describe the design of interfaces within causal domains (those primarily driven by the laws of nature), rather than intentional domains (those driven by human intentions) (Read, Salmon, and Lenne 2012). For HFE practitioners to fully realise the utility of the CWA framework, there is a need for new approaches and guidance for designing beyond interfaces and in different types of domains, using the outputs of CWA. In this article, it is proposed that the values and principles of sociotechnical systems theory can assist to create a theoretically consistent design approach for use with CWA.

Both CWA and sociotechnical systems theory are concerned with the design of sociotechnical systems; being systems that contain both social (human-related) and technical (non-human) aspects that interact to pursue a common goal (Walker et al. 2008). They are both underpinned by the systems perspective and open systems principles. Notably, both aim to design systems that are adaptable in the face of disturbances arising from the external environment. The use of systems-based approaches is especially important in the modern of age of technologically complex, distributed, high-risk domains for which reductionist approaches with assumptions of linearity and rationality are no longer appropriate (Walker et al. 2010; Dekker 2011).

While CWA has been described as a sociotechnical systems approach (e.g. Jenkins et al. 2009; Stanton and McIlroy 2012; Stanton and Bessell 2014), Walker et al. (2008) clarify the distinction between the term sociotechnical systems and sociotechnical systems theory. They note that the former refers to any system of social and technical aspects engaged in goal-directed behaviour, while the latter 'reflects certain specific methods of joint optimisation in order to design

*Corresponding author. Email: gemma.read@monash.edu

Humans as assets

Technology as a tool to assist humans

Promote quality of life

DESIGN PROCESS

Adoption of agreed set of values & purposes

Documentation of how design choices constrain subsequent choices

Provision of resources & support

CONTENT OF THE DESIGNED SYSTEM

Tasks are allocated appropriately between & amongst humans & technology

Useful, meaningful & whole tasks are designed

Boundary locations are appropriate

Design & planning for the transition period

Information provided where action is needed

SYSTEM OUTCOME

Adaptive capacity

Boundaries are managed

Adaptability achieved through flexible structures & mechanisms

Problems controlled at their source

Design driven by good solutions – not fashion

Authority & responsibility allocated appropriately

Needs of the business, users & managers are met

Political debate

System elements are congruent

Adaptability achieved through multifunctionalism

Intimate units & environments are designed

Means for undertaking tasks are flexibly specified

Design is appropriate to the particular context

Multidisciplinary participation & learning

Representation of interconnectedness of system elements

Iteration & planning for ongoing evaluation & re-design

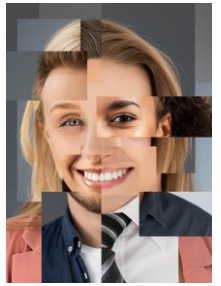
Joint design of social & technical elements

User participation

Adoption of appropriate design process

Constraints are questioned

Respect for individual differences



Responsibility to all stakeholders

Applying sociotechnical theory beyond traditional work design

Received: 28 November 2017 | Revised: 4 August 2018 | Accepted: 6 August 2018

DOI: 10.1002/hfm.20769

RESEARCH ARTICLES

WILEY

A sociotechnical design toolkit for bridging the gap between systems-based analyses and system design

Gemma J. M. Read[1,2] | Paul M. Salmon[1] | Natassia Goode[1] | Michael G. Lenné[2]

[1]Centre for Human Factors and Sociotechnical Systems, Faculty of Arts, Business and Law, University of the Sunshine Coast, Maroochydore, Queensland, Australia

[2]Monash University Accident Research Centre, Monash University, Clayton, Victoria, Australia

Correspondence
Gemma J. M. Read, Centre for Human Factors and Sociotechnical Systems, Faculty of Arts, Business and Law, University of the Sunshine Coast, Locked Bag 4, Maroochydore DC, QLD 4558, Australia.
Email: gread@usc.edu.au

Funding information
Australian Research Council, Grant/Award Numbers: LP100200387, FT140100681, DP120100199, DE180101449, LP110100037; University of the Sunshine Coast

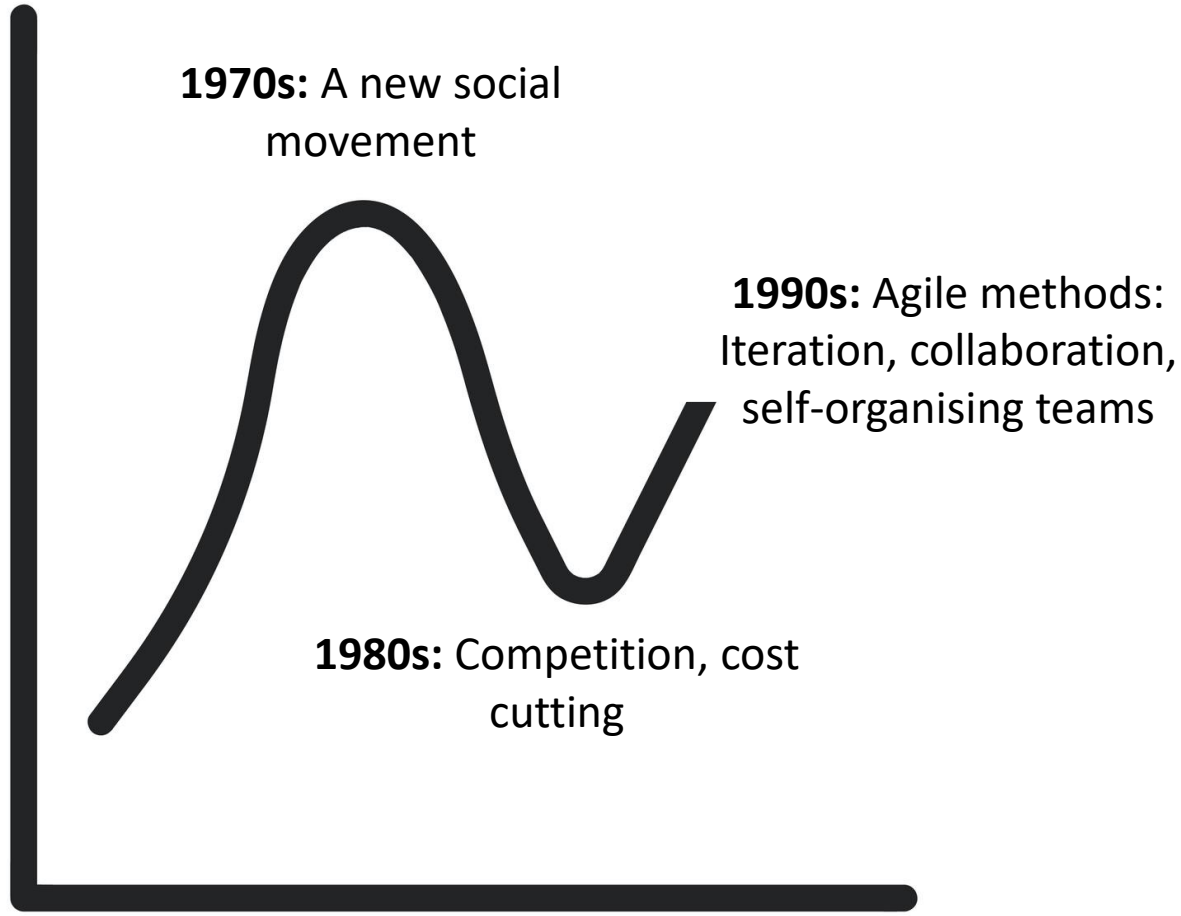
Abstract

Cognitive work analysis (CWA) is a systems-based framework for analyzing complex sociotechnical systems; however, its ability to directly inform design has been questioned. The Cognitive Work Analysis Design Toolkit (CWA-DT) was developed to address this gap. This study describes the development, application, and evolution of the CWA-DT and examines the extent to which it provides a useful approach to design. The evaluation draws upon five case-study applications. In three applications, involving the design of public transport ticketing systems, railway level crossing environments and road intersections, the toolkit was used in conjunction with CWA. In addition, we present two case studies where the CWA-DT was used with other systems-based analysis frameworks (for the design of incident prevention strategies for led outdoor activities and systemic interventions to improve road safety) to illustrate the flexibility of the toolkit. Based on feedback from participants involved in CWA-DT design processes, it is concluded that it provides a useful approach for translating human factors analyses into design concepts. Additional applications and adaptations of the approach are encouraged.

KEYWORDS

cognitive systems engineering, cognitive work analysis, design methods, design strategies, transportation

Rise and fall of STS



The New York Times

A Teen Was Suicidal. ChatGPT Was the Friend He Confided In.

More people are turning to general-purpose chatbots for emotional support. At first, Adam Raine, 16, used ChatGPT for schoolwork, but then he started discussing plans to end his life.

TikTok

This article is more than **2 years old**

TikTok self-harm study results 'every parent's nightmare'

Research suggests algorithm promotes self-harm and eating disorder content within minutes of interest being shown

Dan Milmo *and* Alex Hern

Thu 15 Dec 2022 16.01 AEDT

The researchers set up accounts in the US, UK, Canada and Australia registered with ages of 13. Photograph: Dado Ruvić/Reuters

Does STS remain relevant in the age of AI?

BUSINESS

Tesla factory worker attacked by robot that dug its claws into back and arm: report

By Ariel Zilber

Published Dec. 26, 2023, 2:26 p.m. ET

Amazon scraps 'sexist AI' recruiting tool that showed bias against women

telegraph.co.uk · 2018

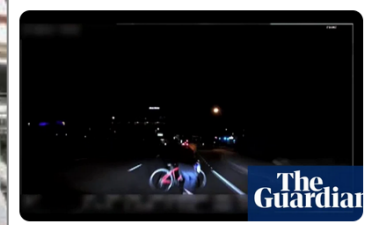
Amazon has scrapped a "sexist" internal tool that used artificial intelligence to sort through job applications.

Fired by Bot at Amazon: 'It's You Against the Machine'

bloomberg.com · 2021

Contract drivers say algorithms terminate them by email—even when they have done nothing wrong.

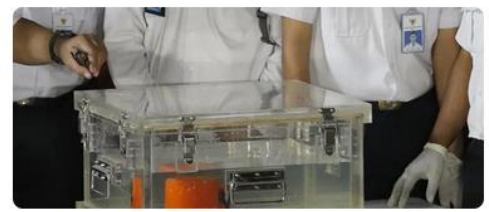
Stephen Normandin spent almost four years racing around Phoenix delivering packages as a contract driver for Amazon.com Inc. Then one day, he received an automated email. The algorithms tracking him had



Uber crash shows 'catastrophic failure' of self-driving technology, experts say

theguardian.com · 2018

Concerns raised about future testing as footage suggests fatal collision in Arizona was failing of system's most basic functions



Pilots struggled against Boeing's 737 MAX control system on doomed Lion Air flight

seattletimes.com · 2018

Data from the fatal Oct. 29 flight that killed 189 people, and from the prior day's flight of the same jet, raises questions about three factors that seem to have contributed to the crash.

https://incidentdatabase.ai/

Current & potential transportation use cases

- Automated transport operations
- Control rooms (e.g. ATC)
- Design and manufacturing
- Maintenance
- Training
- Supervision / task allocation
- Passenger management / security
- Data analytics
- Safety management
- Business management
- Regulatory compliance & legal

Humans as assets

Technology as a tool to assist humans

Promote quality of life

DESIGN PROCESS

Adoption of agreed set of values & purposes

Documentation of how design choices constrain subsequent choices

Provision of resources & support

CONTENT OF THE DESIGNED SYSTEM

Tasks are allocated appropriately between & amongst humans & technology

Useful, meaningful & whole tasks are designed

Boundary locations are appropriate

Design & planning for the transition period

Information provided where action is needed

SYSTEM OUTCOME

Adaptive capacity

Boundaries are managed

Adaptability achieved through flexible structures & mechanisms

Problems controlled at their source

Design driven by good solutions – not fashion

Authority & responsibility allocated appropriately

Needs of the business, users & managers are met

System elements are congruent

Adaptability achieved through multifunctionalism

Political debate

Intimate units & environments are designed

Means for undertaking tasks are flexibly specified

Design is appropriate to the particular context

Multidisciplinary participation & learning

Representation of interconnectedness of system elements

Iteration & planning for ongoing evaluation & re-design

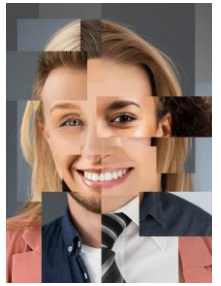
Joint design of social & technical elements

User participation

Adoption of appropriate design process

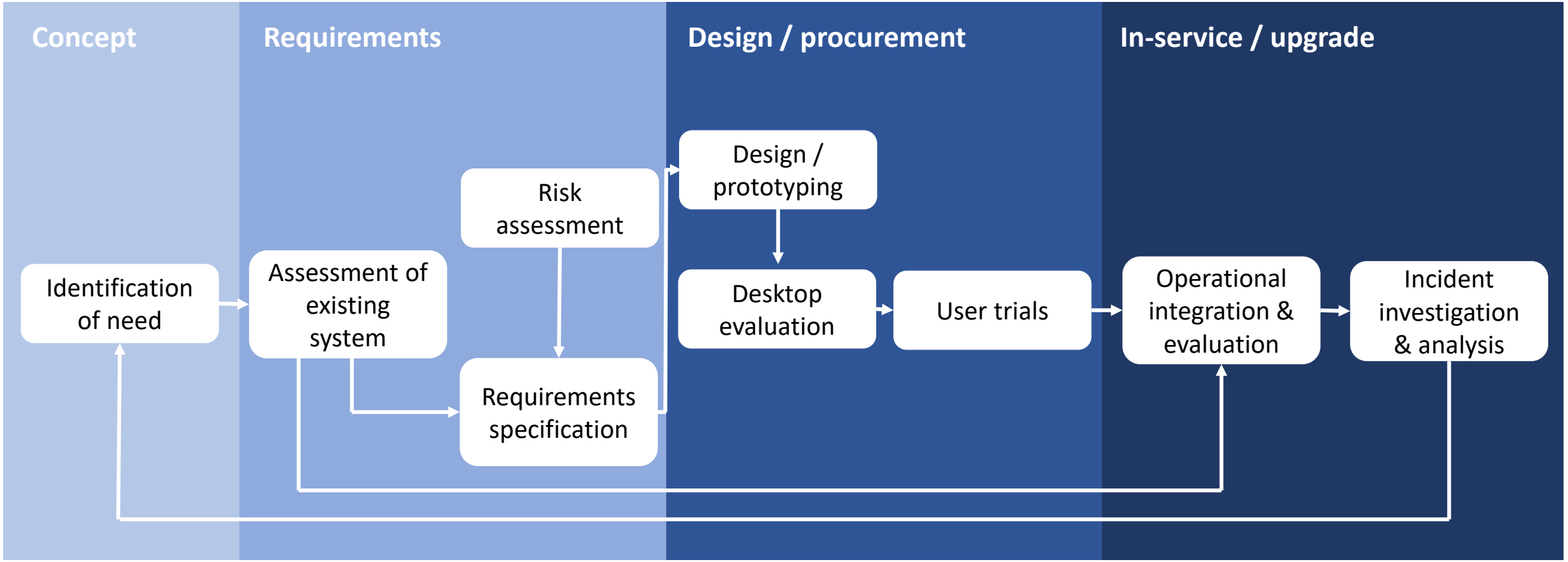
Constraints are questioned

Respect for individual differences

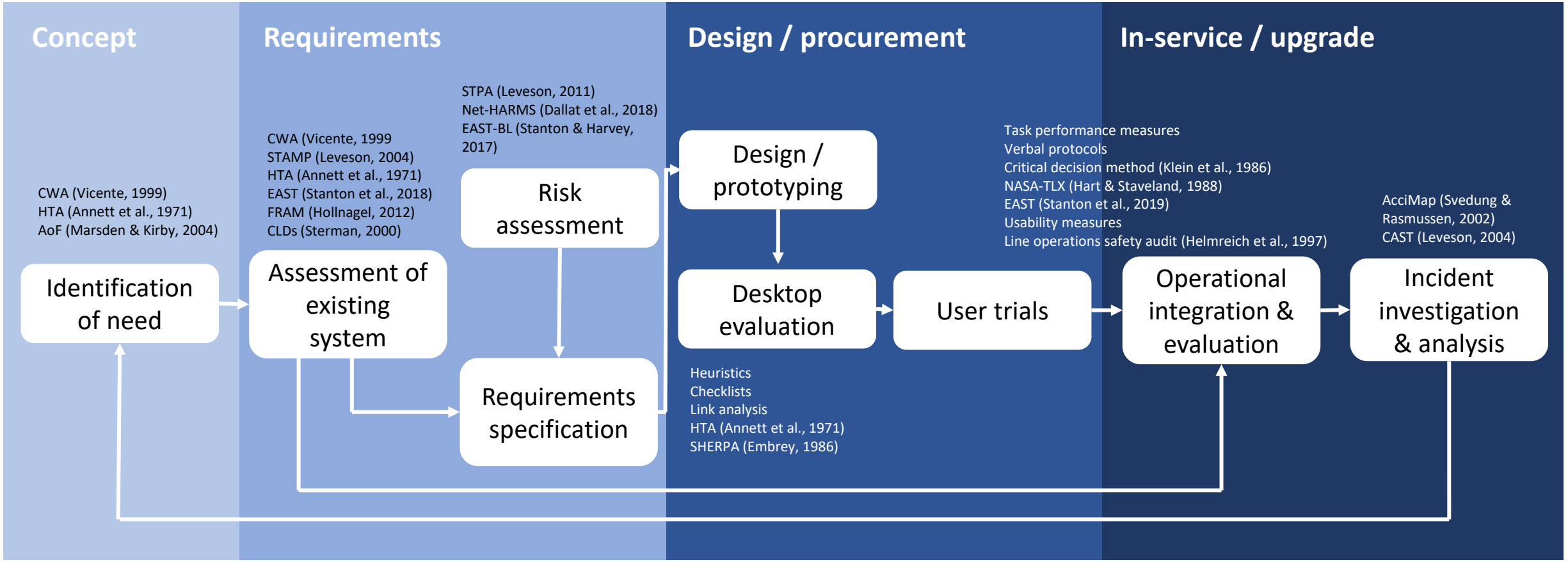


Responsibility to all stakeholders

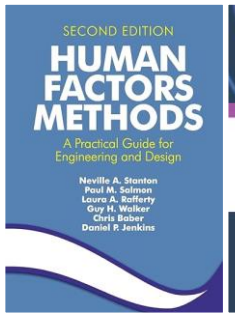
System lifecycle

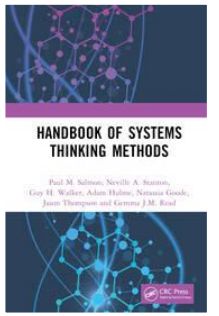


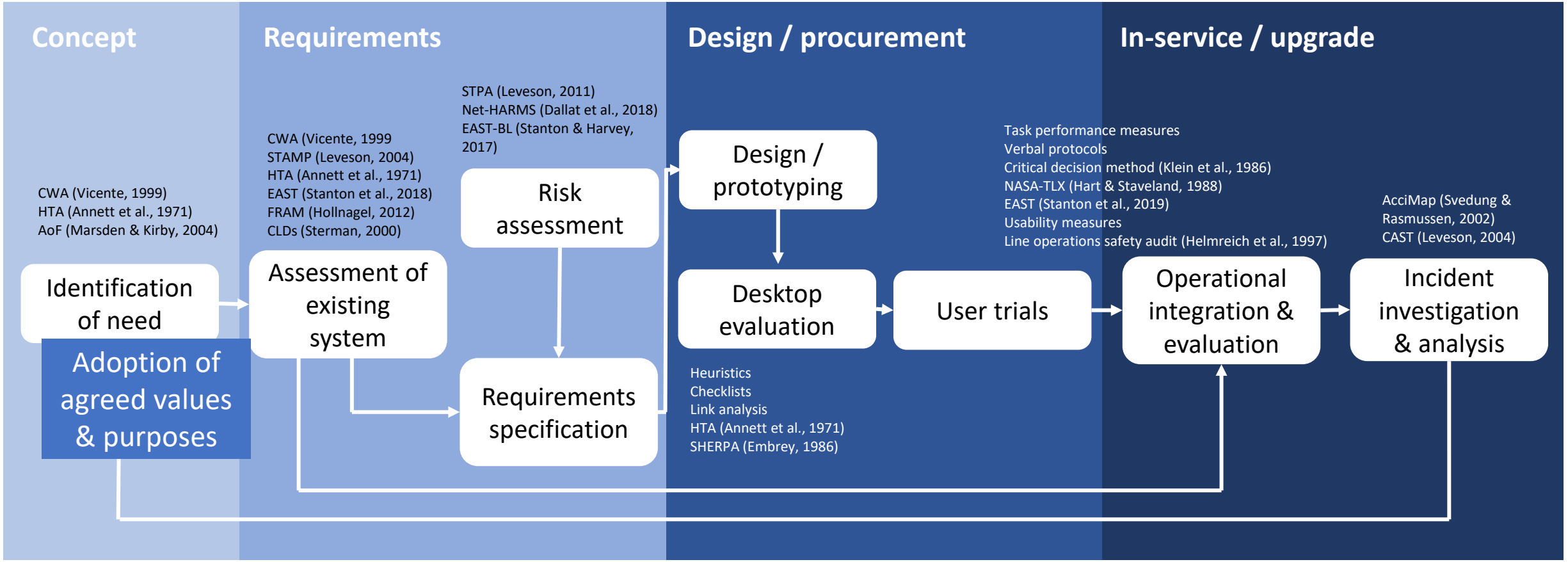
Human Factors toolkit



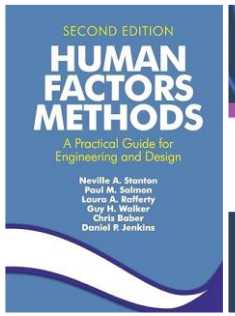
Adapted from Salmon, King, Read, McLean, Shanahan & Devitt (2023)

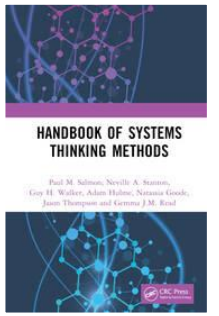




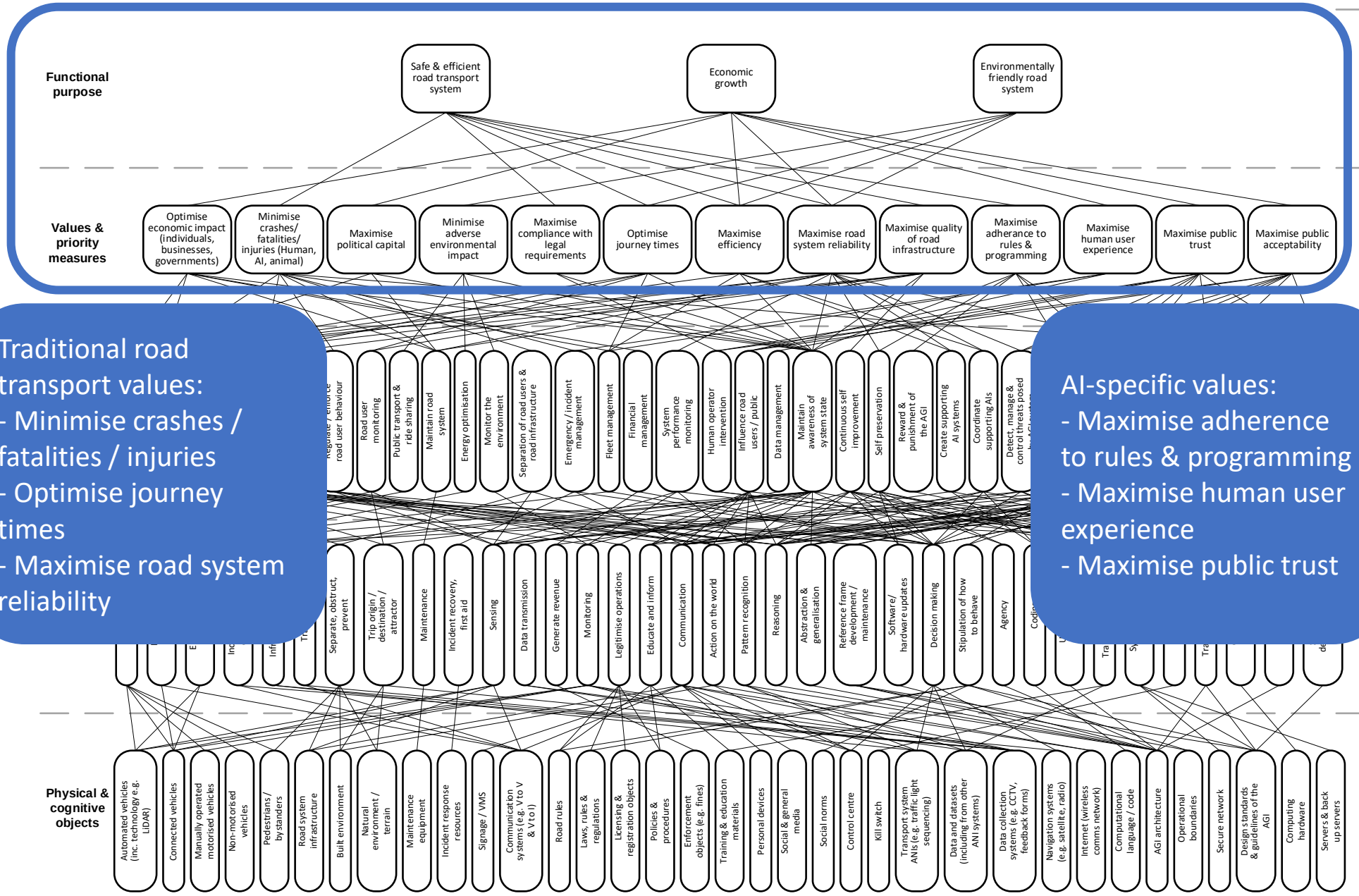


Adapted from Salmon, King, Read, McLean, Shanahan & Devitt (2023)





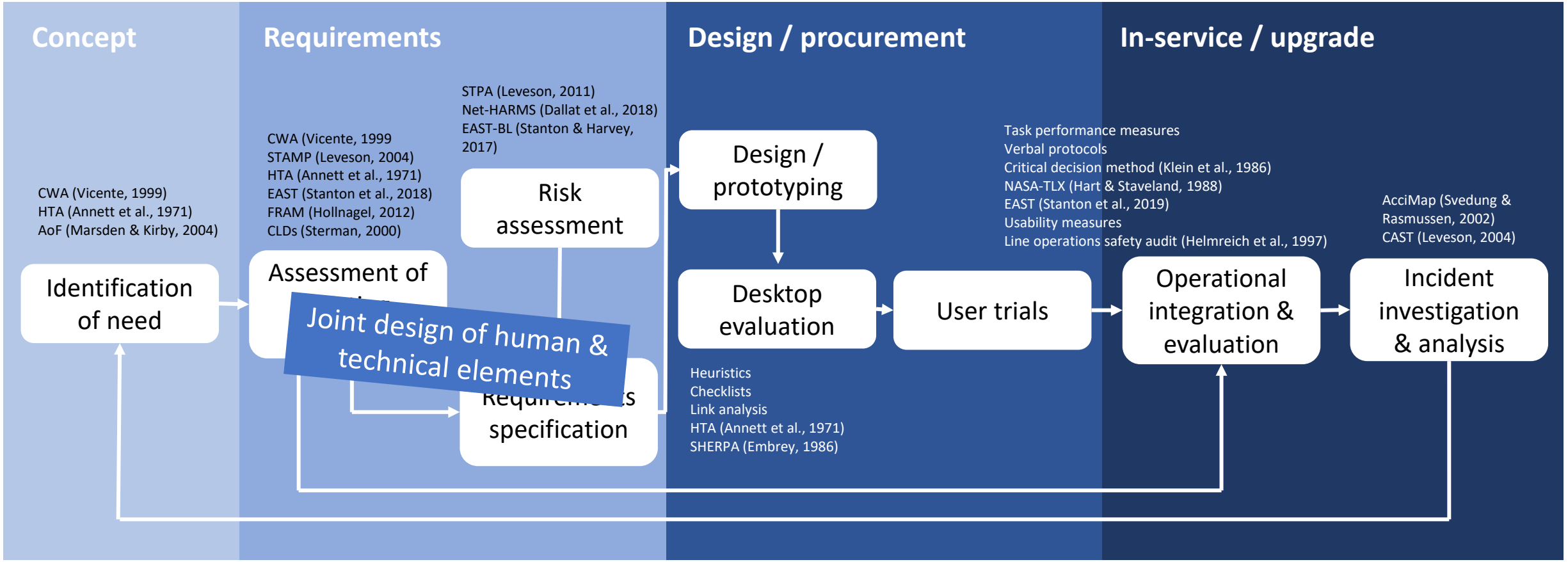
WDA for future AI-based road traffic management system



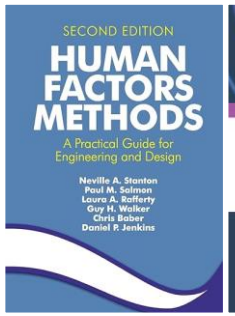
Traditional road transport values:
- Minimise crashes / fatalities / injuries
- Optimise journey times
- Maximise road system reliability

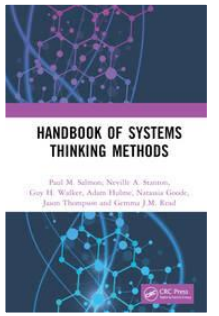
AI-specific values:
- Maximise adherence to rules & programming
- Maximise human user experience
- Maximise public trust

Read et al., 2024

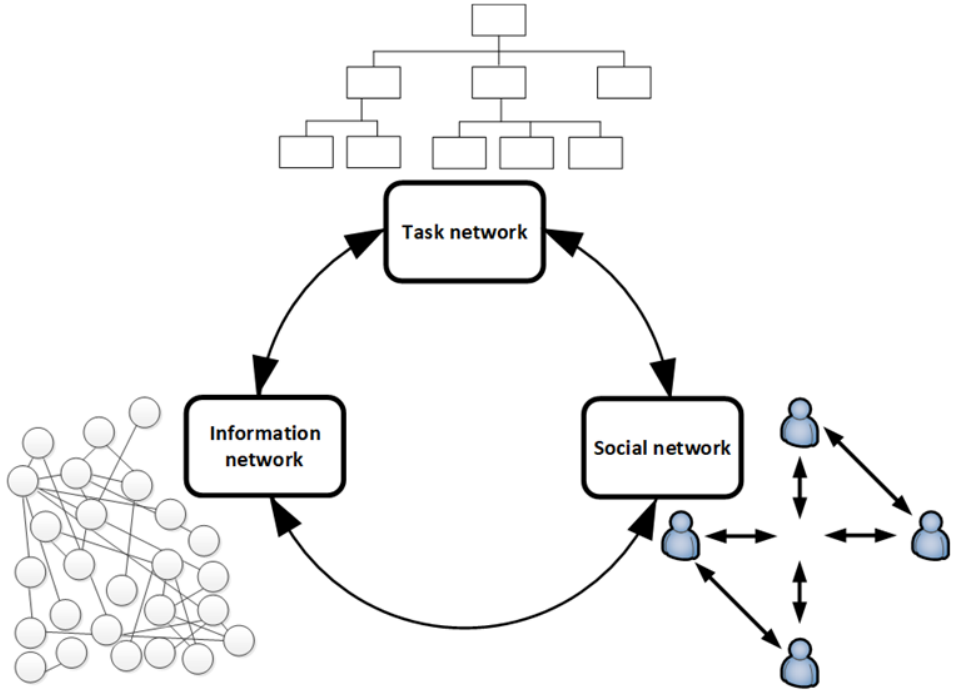


Adapted from Salmon, King, Read, McLean, Shanahan & Devitt (2023)

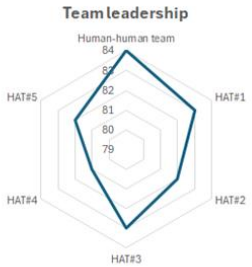


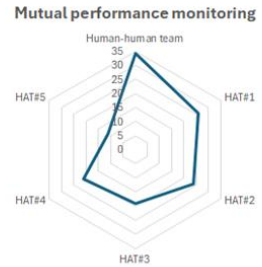


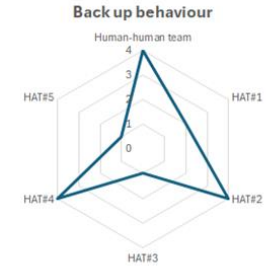
Modelling human-autonomy team configurations in submarine control rooms (Salmon et al., under review)

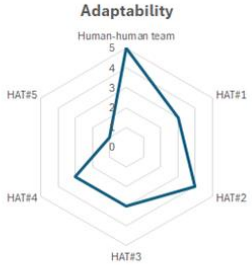


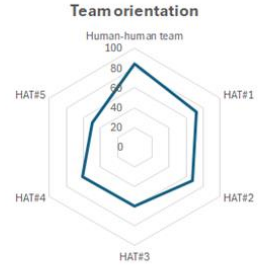
Team configuration	Periscope	Sonar 1	Sonar 2	Track motion analyst 1	Track motion analyst 2	Track manager	Watch leader
HAT#1	Human	Human	AI	Human	Human	Human	Human
HAT#2	Human	Human	Human	Human	AI	Human	Human
HAT#3	Human	Human	AI	Human	AI	Human	Human
HAT#4	Human	Human	Human	Human	Human	AI	Human
HAT#5	Human	AI	AI	AI	AI	Human	Human

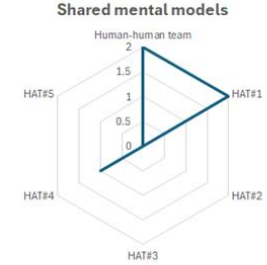


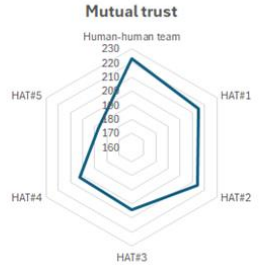


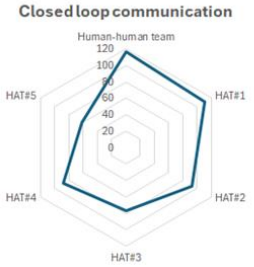




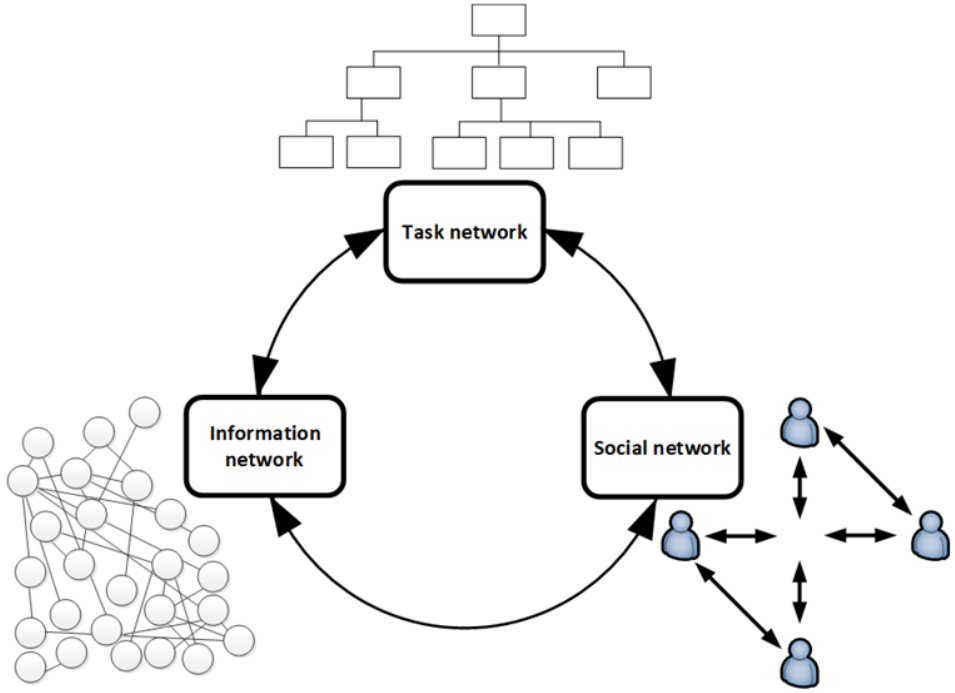




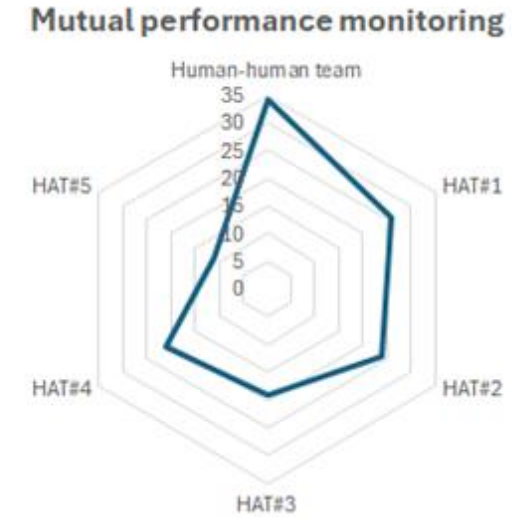


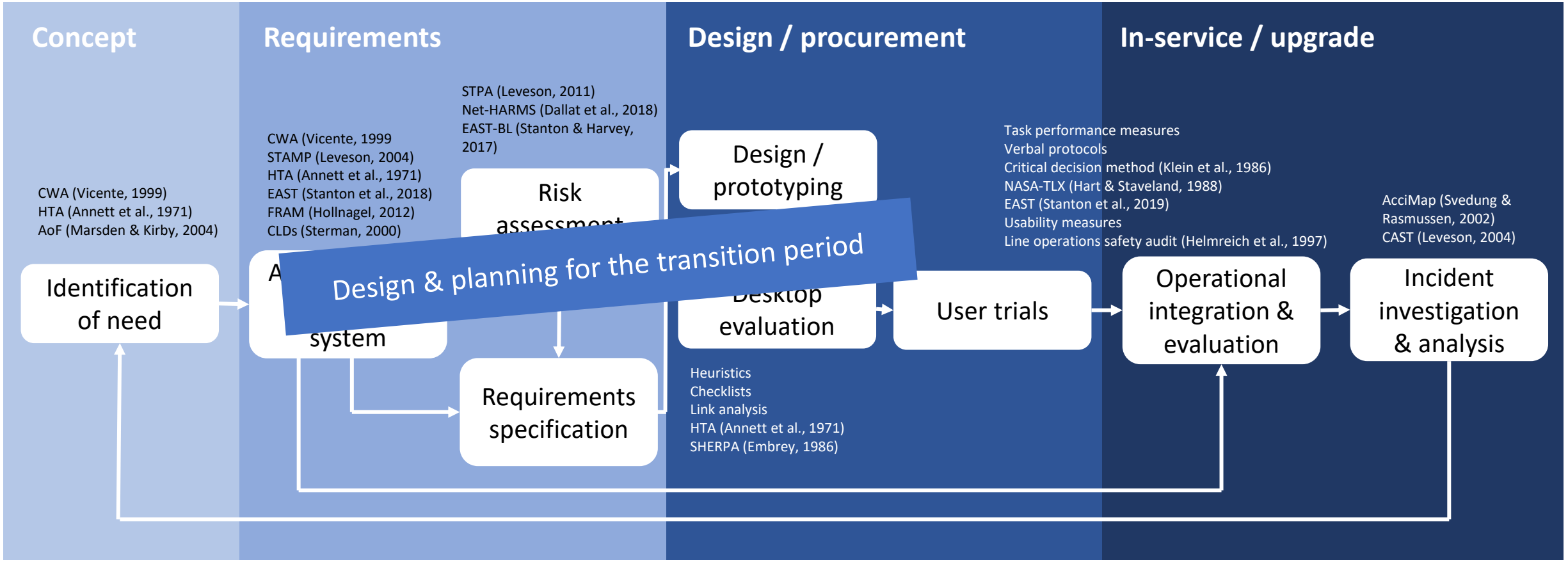


Modelling human-autonomy team configurations in submarine control rooms (Salmon et al., under review)

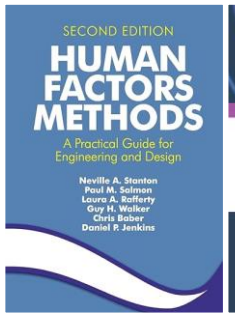


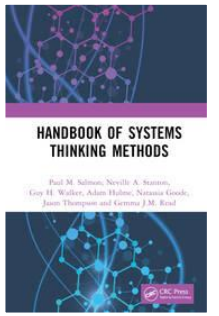
Team configuration	Periscope	Sonar 1	Sonar 2	Track motion analyst 1	Track motion analyst 2	Track manager	Watch leader
HAT#1	Human	Human	AI	Human	Human	Human	Human
HAT#2	Human	Human	Human	Human	AI	Human	Human
HAT#3	Human	Human	AI	Human	AI	Human	Human
HAT#4	Human	Human	Human	Human	Human	AI	Human
HAT#5	Human	AI	AI	AI	AI	Human	Human

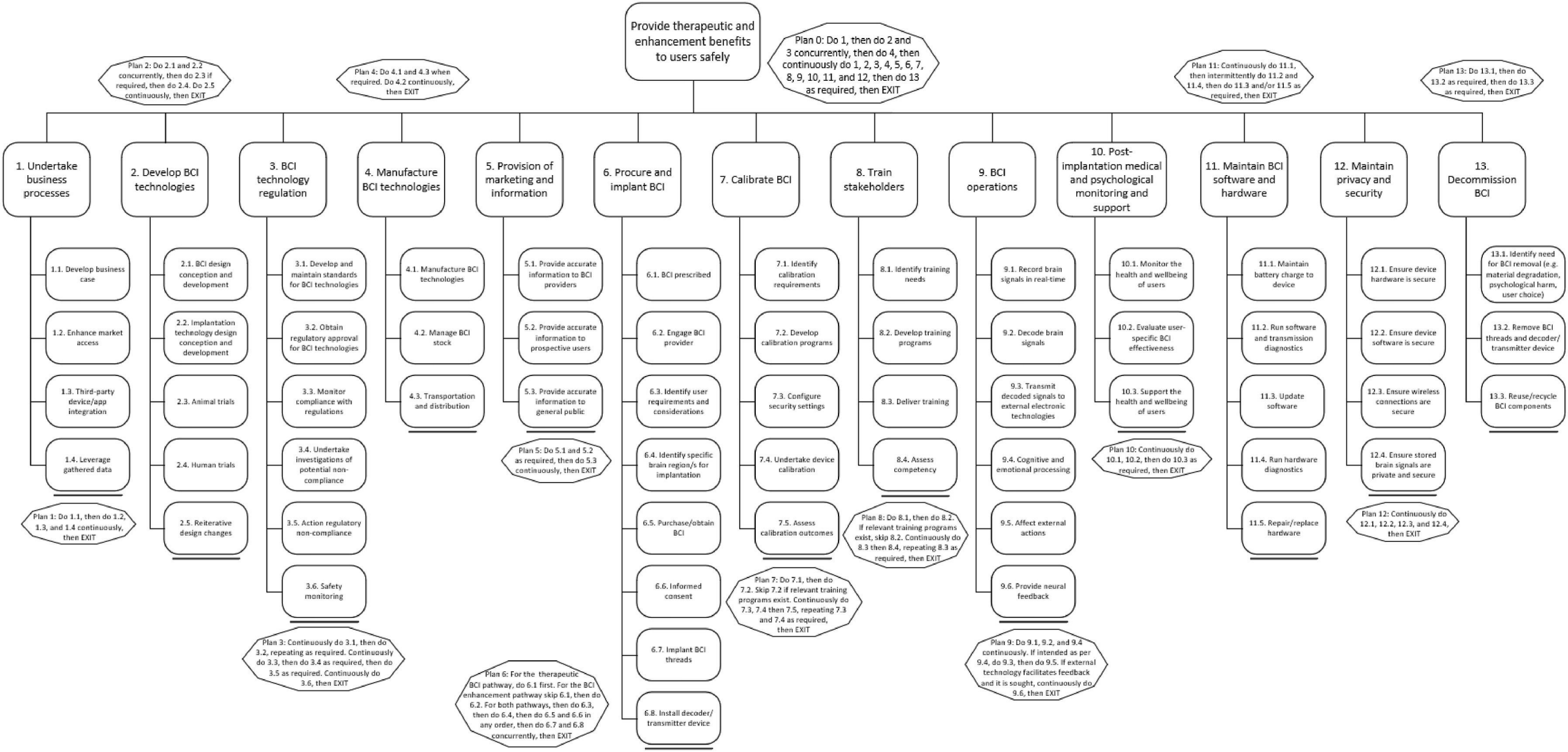




Adapted from Salmon, King, Read, McLean, Shanahan & Devitt (2023)

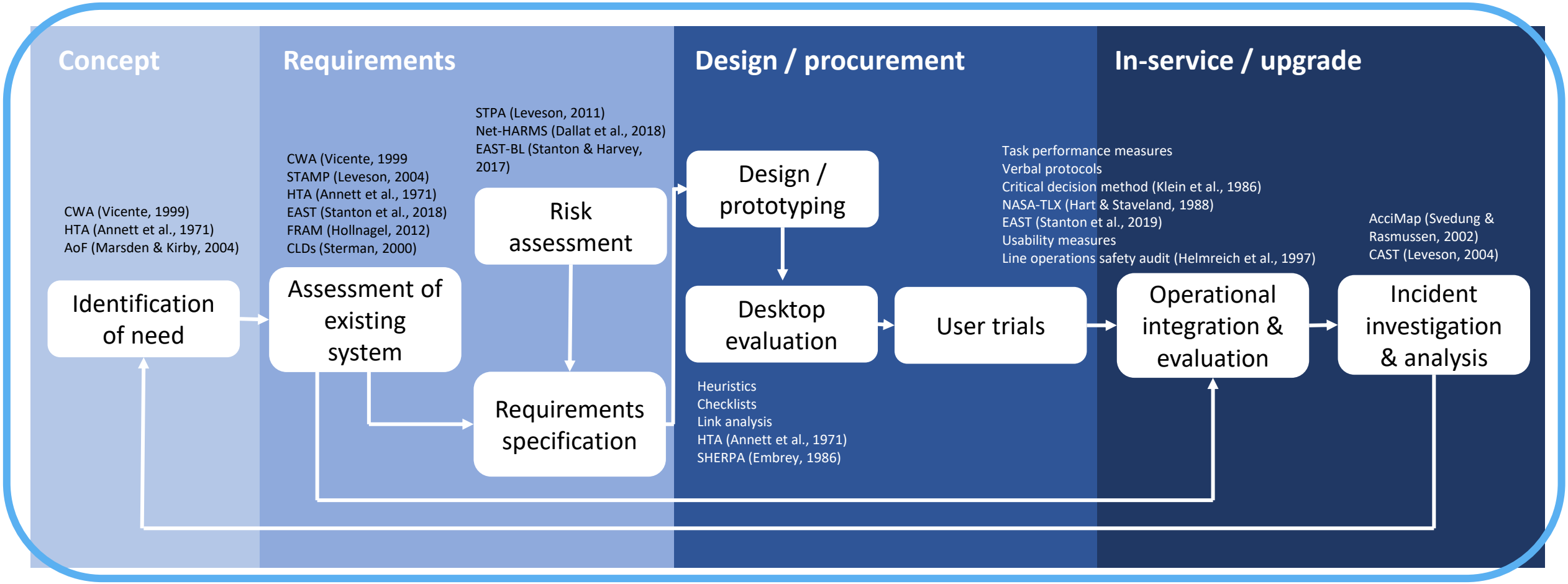






King, Read & Salmon, 2025

Incorporating values across the lifecycle



Are we viewing humans as assets?

Are we viewing technology as a tool to assist humans?

Have we considered quality of life?

Does the system respect individual differences?

Are we meeting our responsibilities to all stakeholders?

Ensuring humans and AI play well together


Credit Adobe Firefly

- Maintain awareness of AI developments
- Be proactive in discussions around AI adoption and facilitate interdisciplinary connections
- Advocate for STS values and principles in design
- Advocate for application of Human Factors methods across the lifecycle
- Collect AI issues & failures
- Promote consideration of future scenarios