This item was submitted to [Loughborough's Research Repository](#) by the author.
Items in Figshare are protected by copyright, with all rights reserved, unless otherwise indicated.

Four studies, two methods and one accident – another look at the reliability and validity of STAMP and Accimap for systemic accident analysis

PLEASE CITE THE PUBLISHED VERSION

<https://doi.org/10.1201/9781315210469-11>

PUBLISHER

© CRC Press.

VERSION

AM (Accepted Manuscript)

PUBLISHER STATEMENT

This work is made available according to the conditions of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International (CC BY-NC-ND 4.0) licence. Full details of this licence are available at:
<https://creativecommons.org/licenses/by-nc-nd/4.0/>

LICENCE

CC BY-NC-ND 4.0

REPOSITORY RECORD

Filho, A.P.G., Gyuchan Thomas Jun, and Patrick Waterson. 2019. "Four Studies, Two Methods and One Accident – Another Look at the Reliability and Validity of STAMP and Accimap for Systemic Accident Analysis". figshare. <https://hdl.handle.net/2134/24056>.

Four Studies, Two Methods, One Accident – Another Look at the Reliability and Validity of Accimap and STAMP for Systemic Accident Analysis

A. P. G. Filho, G.T. Jun & P. Waterson

Human Factors and Complex Systems Research Group, Loughborough Design School, Loughborough University

ABSTRACT: The aim of this study was to assess the validity and reliability of two accident analysis methods. We identified four studies which analysed the same accident (the South Korea Sewol Ferry accident). Studies 1 and 2 used Accimap whilst Studies 3 and 4 applied STAMP. The four studies were compared in terms of analysis procedure taken, granularity of analysis, causal factors identified and recommendations suggested. The results indicate that the reliability between two STAMP studies (61%) is higher than two Accimap studies (31%) in terms of contributing factor identification. It was found that the recommendations made from each study reflect the focus and knowledge of the analyst.

1 INTRODUCTION

The types of methods for accident analysis and investigation that have developed since the 1990's reflect the increasing complexity of socio-technical systems, across a range of domains, including nuclear power production, rail and marine transportation (Waterson et al., 2015; 2017). So-called 'Systemic Accident Analysis' (SAA) draws on sociotechnical systems theory, (e.g., Accimap - Rasmussen, 1997) and control and resilience engineering (e.g., STAMP - Leveson, 2004; FRAM - Hollnagel, 2004) in order to illustrate the diversity of causal factors contributing to accident across different levels of the systems, interactions between them and the role played by external influences such as political, cultural, financial, and technological circumstances (Branford, 2011).

1.1 *Systemic Accident Analysis (SAA): validity and reliability*

Although SAA are underpinned by sociotechnical systems theory, there are significant differences between them in terms of their theoretical assumptions, the type of causal representation and the outputs they produce. The extent to which methods for systemic accident analysis produce outcomes which are valid (e.g., the degree to which the Accimap and STAMP analysis successfully identifies the causes of an accident) and reliable (e.g., the degree to which accident analysts produce similar causal representations) are often viewed as an important criteria for judging their appropriateness for accident analysis (Waterson et al., 2017). According to Underwood and Waterson

(2013), a lack of validation is the key issue which may influence the use of the system approach by practitioners. Additionally, Jacinto and Aspinwal (2004) argue that the ability to learn the right lessons from accidents is dependent upon valid and reliable methods. As such, validation studies are fundamental to any accident analysis method.

According to Katsakiori et al. (2009) methodological problems can partly impact upon reliability. For example, open questions provide the investigator with different information compared to closed questions. Hale et al. (1998) argued that the reliability of results can be affected by not only type of method, but also the background of the analyst such as type of university degree, knowledge and experience. Likewise, accident analysis method evaluation carried out by Gordon et al. (2005) indicated a low overall level of agreement between investigators who had only minimal method training and practice.

The question of validity is closely related to the issue of reliability. If different analysts were to apply Accimap and STAMP to the same accident, would they reach the same conclusions? There have been few attempts to conduct this form of comparative analysis based on an accident analysis (Johnson and Almeida, 2008; Katsakiori et al., 2009). Salmon et al. (2012) provide a comparison between Accimap, HFACS and STAMP based on their analysis of the Mangatepopo gorge incident that was carried out by different experts, but they did not use categories in order to compare the methods (e.g., levels of analysis; causal factors; validity and reliability). Underwood and Waterson (2014) evaluated the ATSB, Accimap and STAMP methods using two criteria:

(1) coverage of systems theory concepts and (2) usage characteristics (e.g., reliability and validity). Both studies, however, did not provide a systematic in-depth evaluation of reliability and validity.

1.2 Study aim

The aim of the present study was to assess the reliability and validity of Accimap and STAMP based on four studies which analyzed the same accident. Two studies applied Accimap (Kee et al., 2016; Lee et al., 2016), while of the other two one applied STAMP (Kim et al., 2016; Kwon, 2016). Three studies were published in peer-reviewed journals (Kee et al., 2016; Kim et al., 2016; Lee et al., 2016) and the final one was a MSc dissertation (Kwon, 2016). In order to compare the studies employed a categorization framework (section 2.1) in order to facilitate the analysis of reliability and validity.

1.3 A brief description of the Accident

On April 16, 2014, Sewol Ferry, the South Korean ship carrying 476 passengers from Incheon to Jeju Island, sank disastrously. The 18-year-old Japanese-built ship was purchased by a company named Chonghaejin, which added two more floors to the ship to hold more passengers, making the ship extremely unstable. During the voyage, when the ship made a sharp turn, it lost its balance and started to list. When the captain Jun Seok Lee communicated with the Vessel Traffic Service (VTS) for help, the captain made questionable decisions such as telling the VTS that the passengers could not evacuate and instructing the passengers to stay on-board. By the time the captain finally told everyone to evacuate, it was too late. As the result of this accident, 304 people, who were mostly high school students, lost their lives in what is considered to be one of the most tragic maritime accidents in the history of South Korea (Lee et al., 2016).

2 METHOD

2.1 A framework for comparing the four studies

In order to systematically compare the four studies, 9 separate categories were used and adapted from Waterson et al., (2017). These covered: (i) the goals and objectives of the study, e.g. providing a systemic account of the factors contributing to the accident; (ii) the theoretical background of the study, e.g. sociotechnical systems theory; (iii) the procedure described in building the Accimap and STAMP, e.g. use of two or more analysts; use of thematic coding or other qualitative data analysis methods; (iv) the outcomes from the analysis, e.g. type of Accimap; (v) comparisons with other models; Acimmap; (vi) levels of analysis, i.e., number and type of levels

used; (vii) casual factors; (viii) changes to the standard Accimap and STAMP, i.e. major/minor deviations from the format of the Accimaps described in Rasmussen and Svedung (2000); and (ix) other details, e.g. source of the data for accident analysis.

2.2 Reliability and validity analysis

The validity and reliability of the four studies were evaluated qualitatively and quantitatively. It was measured that: (1). the number of common factors described between two studies was identified; (2). the number of common factors was divided by the total number of factors described by two different accident analysis methods when convergent validity was assessed or by the same accident analysis methods when reliability was assessed and expressed as a percentage value (Jacinto and Aspinwall, 2004). For example, in order to assess convergent validity, the number of common factors described in Studies 1 (Accimap) and 3 (STAMP) was divided by the total of factors described in both studies. Similarly, in order to assess reliability the number of common factors (within the same level) described in Studies 1 (Accimap) and 2 (Accimap) was divided by the total of factors described in both studies.

3 FINDINGS

Table 1 is summary of the results of applying the framework to the four studies. In subsequent Sections of the paper, we refer to the studies with reference to their number in Table 1 (e.g., Study 1 is Lee et al., 2016).

3.1 Goals, objectives and theoretical background

The four studies have the same overall goal which was to apply a system approach to analyze the South Korea Ferry accident and identify the range of contributing factors in causing the accident to happen. Studies 1 and 2 applied the Accimap to analyses systematically the accident and identify the contributing factors as well as interactions between them. Studies 3 and 4 applied the STAMP method to the entire maritime transportation socio-technical system in order to provide a broader view of factors contributing to the accident. Socio-technical systems theory was adopted as theoretical background by all studies, whereas Rasmussen's (1997) framework was also explicitly adopted as a theoretical base by studies 1, 2 and 3.

Table 1 – Applying the framework to the four studies

	Accimap		STAMP	
	Study 1 (Lee et al., 2016)	Study 2 (Kee et al., 2016)	Study 3 (Kim et al., 2016)	Study 4 (Kwon, 2016)
Goals and objectives	To Apply the Accimap methodology to identify the contributing causes as well as interactions between them	To analyze systematically the South Korea Ferry accident using the Accimap	To illustrate the utility of applying the STAMP Model to the maritime transportation domain	To show the application of CAST to investigate the Sewol-Ho Ferry Accident
Theoretical background	Sociotechnical systems theory	Sociotechnical systems theory	Sociotechnical systems theory	Sociotechnical systems theory
Procedure	The accident analysis was carried out by study authors but there no evidence that Accimap was validated by analysts	Two authors of the study conducted an AcciMap analysis of the accident. This was internally validated by the remaining co-authors and externally validated by the senior operation manager in a shipping company	The accident analysis was carried out by study authors but there no evidence that the STAMP was reviewed by analysts	The accident analysis was carried out by study authors but there no evidence that the STAMP was reviewed by analysts
Outcomes	One Accimap was produced	Two Accimap were produced	One hierarchical safety control structure was produced	Two hierarchical safety control structure were produced
Comparison with other models	None	None	None	None
Levels of analysis	6 main layers	6 main layers	6 levels of the hierarchical safety control structure	10 levels of the hierarchical safety control structure
Causal factors	28 factors	49 factors	28 factors	379 factors
Changes to standard Accimap/Stamp	No major changes	No major changes	No major changes	No major changes

3.2 Procedure and outcomes

In Study 2, two method experts were involved in creating an initial Accimap, which was validated by another expert and analysis of the accident was conducted by two authors of the Study 2, this was internally validated by the remaining co-authors and externally validated by a senior operation manager in a shipping company. The rest of the studies did not provide the validation procedures.

The Accimap framework developed by Rasmussen and Svedung (2000) for the analysis of the Zeebrugge Ferry accident was used in Study 1, whereas Study 2 adopted the Accimap structure used in Svedung and Rasmussen (2002) and Branford (2011). The STAMP (CAST, Leveson, 2004) nine steps were used in Studies 3 and 4.

In Study 1 and 3, one representation, e.g. one Accimap and one hierarchical safety control structure, was used whereas Study 2 and 4 used two separate representations dividing the Accimap or the STAMP control structure into two parts of the capsizing of the ferry and the rescue operation.

3.3 Levels of analysis, causal factors and other details

Both Study 1 and 2 analyzed the six levels as described by Rasmussen (1997), but re-labelled levels six to *outcome*. Study 3 and 4 analyzed the levels of the hierarchical safety control structure including the development and operation of ships in South Korea.

Study 3 analyzed six levels (e.g., Ministry of Oceans and Fisheries; Korean Coast Guard; Korean Register of Shipping; and industry association; Ship-owning company; Master and Crew), whereas Study 4 analyzed four additional levels (e.g., Loading Service Company (LOSC); Life Rafts Inspection Company (LIC); Lashing Service Company (LASC); Passengers).

Study 2 and 4 divided the accident analysis into two separate representations (one for capsizing and the other for poor rescue), while Study 1 and 3 included both in one representation. Consequently, Study 2 and 4 have a higher level of detail, identifying 49 and 379 contributing factors respectively. Study 2 and 4 identified 13 and 28 respectively. In terms of the data source, all the four studies were based on the original investigation reports of Korean governmental agencies (e.g., Korean Maritime Safety Tribunal) and major South Korean and foreign newspapers and news magazines.

3.4 Comparison between STAMP and Accimap outcomes

Study 2 (Accimap) and Study 3 (STAMP) were selected to check how many common contributing factors these two different methods identified. Study 2 and 3 were selected because both studies identified a very similar number of contributing factors (30 and 28).

As shown in Table 2, six common factors were identified out of fifty two total number of factors identified by both methods (12%). In other words, 20% (Accimap) and 21% (STAMP) of factors identified by each method are common factors.

Table 2 – Comparison between STAMP and Accimap outputs – contributing factors

	Accimap	STAMP	Total factors
	Study 2	Study 3	
Identified factors	30	28	52
Common Factors	6	6	6
%	20	21	12

3.5 Comparison between contributing factors identified from the same method

Table 3 shows the comparison of outcomes between the same methods. Study 1 and 2 had common factors at the level of shipping company and master/crew. Study 2, which analysed at a higher level of detail than Study 1, identified factors at the most of categories (11/13) while Study 1 identified factors at

5 categories (5/13). For example, Study 1 described a factor “Lack of a central authoritative or oversight” at the level of Government, whereas Study 2 described a similar factor more specifically “Ministry of Oceans and Fisheries: Lack of Technical Guidelines for Ship Re-modelling.” Additionally, Study 1 focused on legislation, e.g., “in 2008, under the Lee administration, the maximum allowable age for a passenger ship went to form 20 years to 30 years” and Study 2 more broadly described various factors on “Culture of Society”, “Pro-Business Policy” and “Low Profitability”.

Study 4 had four additional levels of the hierarchical safety control structure compared to Study 3 including various service providers and passengers. Study 4 divided the accident analysis into two parts (capsizing and poor rescue) like Study 2 and analysed a very high level of details. For example, at the level of Master and Crew, Study 4 analyzed each member, including the Captain, First Mate, Second Mate, Third Mate, Helmsman, and crewmembers in the engine room, whereas Study 3 analyzed Master and Crew as a whole. As a result, Study 3 described 11 factors at the level of Master and Crew, whereas Study 4 described 120. Another distinguishable difference is that Study 3 did not describe factors on the poor rescue operation at the level of Korean Coast Guard (KCG), but Study 4 described 105 factors. Four common factors were identified between outcomes from two Accimap studies and seventeen common factors were identified from two STAMP studies. It means that 31 % of factors identified in Study 1 were also identified by Study 2 (higher level of details) and 61% of factors identified in Study 3 were also identified by Study 4 (higher level of details).

3.6 Study recommendations

In Study 1 the recommendations focuses on legislative and regulatory changes (e.g. the national government and local governments need to strengthen enforcement of rules and procedures). Study 2 focuses on South Korean culture and their recommendations involve wide-scale changes to the way in which socio-political and economic pressures should be relaxed in order to promote safety. Study 3 recommends improvement to the hierarchical safety control structure that ensures safe development and operation of passenger ships in South Korea as a whole. Finally, Study 4 produced more comprehensive and detailed recommendations (41 in total), covering each level of the hierarchical safety control structure that ensures safe development and operation of passenger ships in South Korea (e.g., ROKCG should review and enhance training on rescue operation and communication system).

Table 3 – Comparison between factors from the same method

	S1	S2	Common	S3	S4	Common
	Factors		Factors	Factors		factors
Ministry of Oceans and Fisheries (MOF)	0	4	0	3	10	3
Korean Register of Shipping (KRS)	1	2	0	1	13	1
Korean Coast Guard (KCG)	2	3	0	3	113	2
Korean Shipping Association (KSA)	2	2	0	3	32	2
Shipping Company	4	3	1	7	51	3
Master and Crew (MC)	4	4	3	11	120	6
Ministry of Security and Public Administration (MSPA)	0	3	0	0	0	0
Central Disaster and Safety Countermeasures Headquarter (CDSCH)	0	2	0	0	0	0
Regional Rescue Centre (RRC)	0	2	0	0	0	0
Incheon Regional Maritime Affairs and Port Administration (IRMAPA)	0	1	0	0	0	0
Culture of Society (CS)	0	4	0	0	0	0
Service Company	0	0	0	0	27	0
Passenger	0	0	0	0	13	0
Total	13	30	4	28	379	17
Overlapping ratio (common factors/factors in study 1 or 3)			31%			61%

S1, 2, 3 and 4 = Study 1, 2, 3 and 4

4 DISCUSSION

4.1 Accident analysis method characteristics

An important goal of accident analysis is to provide an accurate assessment of the most important contributory factors that triggered human error or mistakes to be made. Lundberg et al. (2009) argue, however, that methods may differ on how they capture this ‘reality’. One model may focus on factor X, while another may highlight factor Y. Some meth-

ods, because of the characteristics may miss factors that others deem important. In our analysis of the four studies, STAMP (studies 3 and 4) focused on providing an accurate hierarchical safety control structure, whilst Accimap (studies 1 and 2) highlighted general aspects of the system (e.g., legislation, culture of society, pro-business policy).

Additionally, different facets of the accident were explored by the four studies, even when the same accident method was used (e.g., study 1 explored the need for improved legislation, whereas study 2 explored wider societal factors).

Finally, STAMP method has another characteristic that is the inclusion of the context in which decisions were made as well an emphasis on cognitive aspects leading up to the accident (e.g., flawed mental models flaws). This characteristic helps to better understand why something went wrong and introduces more factors in accident analysis. For instance, according to Study 3 (STAMP) the *lack of experience and training* (context in which decision was made) of the Master and Crew *led them to underestimate of outcome of the sudden turn* (mental model flaw). In turn, Accimap has an important feature that is the linkages of factors within and between levels, which help to understand inter-relationships between them and causality across the system levels. This feature helped to understand, for example, the inter-relationships between following factors across the marine transportation system in South Korean that led to ferry accident presented in Accimap produced by Study 2: *inappropriate patronage practices* (level 1) $\Rightarrow$ *lack of independence in the position of marine operating inspector* (level 2) $\Rightarrow$ *poor safety operation check* (level 4) $\Rightarrow$ *lack of proper securing* (level 5) *cargo shift* (level 5) $\Rightarrow$ *capsized ferry* (level 6).

4.2 The role of the analyst

Accimap and STAMP afford considerable flexibility in terms of the way in which the analyst can choose to use, interpret and decide upon the outputs from using either method. In many respects this can be an advantage (e.g., providing greater freedom in ‘exploring’ underlying factors within the accident, as compared to deriving a set of firm conclusions) and a disadvantage (e.g., providing a threat to reliability/validity). The profile of analyst (e.g., background, experience, world view) can influence the focus on and produce variation in investigation findings (Underwood and Waterson, 2013). Hollnagel and Speziali (2008) similarly argue that is simply not possible to begin an investigation with a completely open mind, just as it is not possible passively to ‘see’ what is there. From these points of view, it might be explained the different focus on given by each study and the variation in findings in all of them, even when the same method was used. The backgrounds and profiles of the authors across the studies may

explain, for example, why study 2 was broken into two parts the different objectives and outcomes used and derived across all four studies.

4.3 Validity and reliability

Comparing method with others is common in validation and reliability studies, but care must be taken when drawing conclusions. The analysis carried in this study was of a qualitative, rather than a quantitative, nature as it involved subjective judgment in deciding which factors were common, given the fact that the four studies used different nomenclatures in their factors.

Although it is difficult to establish what percentage should be considered a “good result”, for there are no reference values of other identical studies against which to compare. The results in this study indicate that the reliability between two STAMP-based studies (61% overlap) is higher than the reliability of two Accimap studies (31% overlap). Considering 20-21% overlap between Study 2 (Accimap) and Study 3 (STAMP), it shows that use of the same method is more likely to identify the same contributing factors than use of the different method. Given qualitative characteristics and lack of taxonomy of both Accimap and STAMP (more flexibility on Accimap), as has already been mentioned, the result was understandable. Low reliability of Accimap and STAMP has been reported by others studies (Hollnagel and Speziali, 2008; Underwood and Waterson, 2014; Waterson et al., 2017), but this study more systematically compared the difference.

Despite relatively low validity and reliability presented by all four studies, both Accimap and STAMP enabled the analysts explore different possibilities to explain the Korean ferry accident. On other hand, if a taxonomy were constructed to both of them, how has been suggested by many researchers (Salmon et al., 2012; Waterson et al., 2017), it could improve the output from them, yet the freedom of analyst to explore the accident would be lost. From these four studies, it might be more important explore accident than apply or be bound by strict criteria such as validity and reliability. This is important and complex trade-off that warrants further investigation (Waterson et al., 2014).

4.4 Conclusion and future work

All of the studies aimed to provide a systemic account of factors which contribute towards South Korea Sewol ferry accidents. Part of this involved emphasizing that accidents do not come about as a result of single failures, but are attributable to a range of causal factors distributed across various part of the larger system. Despite the popularity of this view amongst the research community, single factor explanations often prevail. A key advantage of SAA

methods such as STAMP and Accimap is that they avoid these types of ‘reductionist’ explanations and support systems thinking (Dekker, 2014). Our work in this study has raised some questions however, about how simple it is to use criteria such as reliability and validity with these types of methods. Accident analysis and investigation normally places a high value on reliability and validity (e.g., Katsakiori et al., 2009; Ryan, 2015). The strength of established methods such as Fault and Event Tree Analysis for example, is that they are often perceived as reliable and valid. SAA methods by contrast, may be more usefully thought as providing additional coverage of organizational ‘big picture’ information (Branford, 2011). Information that may be missing from traditional methods. Rather than seeing SAA methods as a replacement for traditional methods, it might be worthwhile to explore their potential to be combined with methods such as FTA. Part of our current work involves a set of studies examining how this might work in domains such as healthcare (e.g., using STAMP/Accimap as an ‘add-on’ to popular methods such as Root Cause Analysis).

5 REFERENCES

- Branford, K. (2011), Seeing the big picture of mishaps: applying the Accimap approach to analyse system accidents. *Aviation Psychology and Applied Human Factors*, 1, 1, 31-37.
- Dekker, S.W.A. 2014. *The Field Guide to Understanding Human Error* (3rd Edition). Farnham: Ashgate.
- Gordon, R., Flin, R. and Mearns, K., 2005. Designing and evaluating a human factors investigation tool (HFIT) for accident analysis. *Safety Science*, 43, 147-171.
- Hale, A., Heming, B., Smit, K., Rodenburg, F. and van Leeuwen, N., 1998. Evaluating safety in the management of maintenance activities in the chemical process industry. *Safety Science*, 28, 21-44.
- Hollnagel, E., 2004. Barriers and Accident Prevention. Ashgate, Farnham.
- Hollnagel, E. and Speziali, J., 2008. Study on Developments in Accident Investigation Methods: A Survey of the “State-of-the-Art”. *SKI Report 2008:50*.
- Jacinto, C. and Aspinwal, E., 2004. WAIT (Part III) - Preliminary validation studies. *Safety Science Monitor*, 8, 19-29.
- Johnson, C.W. and de Almeida, I.M., (2008). An investigation into the loss of the Brazilian space programme's launch vehicle VLS-1 V03. *Safety Science*, 46, 1, 38-53.
- Katsakiori, P., Sakellaropoulos, G. and Manatakis, E., 2009. Towards an evaluation of accident investigation methods in terms of their alignment with accident causation models. *Safety Science*, 47, 1007-1025.
- Kee, D., Jun, G. T., Waterson, P.E. and Haslam, R., 2016. A systemic analysis of South Korea Sewol Ferry accident - Striking a balance between learning and accountability. *Applied Ergonomics*, 1-13.
- Kim, T., Nazir, S. and Øvergård, K. I., 2016. A STAMP-based causal analysis of the Korean Sewol ferry accident. *Safety Science*, 83, 93-101.
- Kwon, Y., 2016. System Theoretic Safety Analysis of the Sewol-Ho Ferry Accident in South Korea. Master of Science in Engineering and Management at the Massachusetts Institute of Technology.

- Lee, S., Moh, Y. B., Tabibzadeh, M. and Meshkati, N., 2016. Applying the Accimap methodology to investigate the tragic Sewol Ferry accident in South Korea. *Applied Ergonomics*, 1-9.
- Leveson, N., 2004. A new accident model for engineering safer systems. *Safety Science*, 42, 237-270.
- Lundberg, J., Rollenhagen, C. and Hollnagel, E., 2009. What-You-Look-For-Is-What-You-Find – The consequences of underlying accident models in eight accident investigation manuals. *Safety Science*, 47, 1297-1311.
- Rasmussen, J., 1997. Risk management in a dynamic society: a modeling problem. *Safety Science*, 27, 2, 183-213.
- Rasmussen, J., Svedung, I., 2000. Proactive Risk Management in a Dynamic Society, first ed. Raddningsverket, Risk and Environmental Department. Swedish Rescue Services Agency, Karlstad, Sweden.
- Ryan, B. 2015. Incident reporting and analysis. In J.R. Wilson and S. Sharples (Eds.), *Evaluation of Human Work* (4th Edition). Boca Raton: CRC Press.
- Salmon, P., Cornelissen, M. and Trotter, M.J., 2012. Systems based analysis methods: a comparison of Accimap, HFACS and STAMP. *Safety Science*, 50, 1158-1170.
- Svedung, I., Rasmussen, J., 2002. Graphic representation of accident scenarios: mapping system structure and the causation of accidents. *Safety Science*, 40, 397-417.
- Underwood, P. and Waterson, P.E., 2013. Systemic accident analysis: Examining the gap between research and practice. *Accident Analysis and Prevention*, 55, 154-164.
- Underwood, P. and Waterson, P.E., 2014. Systems thinking, the Swiss Cheese Model and accident analysis: A comparative systemic analysis of the Grayrigg train derailment using the ATSB, AcciMap and STAMP models. *Accident Analysis and Prevention*, 68, 75-94.
- Waterson, P.E., Clegg, C.W. and Robinson, M. 2014. Trade-offs between reliability, validity and utility in the development of human factors methods. In O. Broberg, N. Fallentin, P. Hasle, P.L. Jensen, A. Kabel, M.E. Larsen, T. Weller (Editors), *Human Factors in Organizational Design and Management XI*. Santa Monica: IEA Press
- Waterson, P.E., Robertson, M.M., Cooke, N.J. Militello, L., Roth, E. and Stanton, N.A. 2015, Defining the methodological challenges and opportunities for an effective science of sociotechnical systems and safety. *Ergonomics*, 58, 650-8.
- Waterson, P.E., Jenkins, D. P., Salmon, P. M. and Underwood, P., 2017. 'Remixing Rasmussen': The evolution of Accimaps within systemic analysis. *Applied Ergonomics*, 59, Part B, 483-503. .