ISSN 2230-5734

JournalofSAR.com

VOLUME 8
ISSUE 2

Supported by:

ICPEM

EXCELLENCE IN DISASTER RESEARCH

JOURNAL of SEARCH+RESCUE

Volume 8, Issue 2, December 2025
Pages 55-122
ISSN 2230-5734
https://doi.org/10.61618/DHXM4316

The Journal of Search and Rescue (JSAR) is an open access peer-reviewed electronic journal for the collation and distribution of original scholarly material on search and rescue (SAR).

It is being supported by the in-kind work and contributions of the Editorial Board. There is still the need for a dedicated journal serving those with a direct interest in all disciplines of search and rescue including: rope rescue, water (flat, swift and marine), ice rescue, wilderness search and rescue, structural collapse rescue, trench collapse rescue, cave rescue, dive rescue, motor vehicle extrication, canine search, technical animal rescue, air rescue, search theory, search management, and mines rescue. JSAR exists to fulfil that need.

Article submissions from these and other SAR disciplines are welcome. Launching this journal on the internet offers a relatively cost-effective means of sharing this invaluable content. It affords the prompt publication of articles and the dissemination of information to those with an interest in SAR.

JSAR will provide a forum for the publication of original research, reviews and commentaries which will consolidate and expand the theoretical and professional basis of the area. The Journal is interested in theoretical, strategic, tactical, operational and technical matters.

Advertising within JSAR will be considered in the future to ensure sustainable funding is available to enhance and continue the work of the journal. The publication of an article in the Journal of Search and Rescue does not necessarily imply that JSAR or its Editorial Board accepts or endorses the views or opinions expressed in it.

Cover Image
A) *Subject: Search Dog Layla and Chloe Creighton*, K9 Search & Rescue, 2024 *Photographer:* Ryan Gray

Volume 8, Issue 2, December 2025
Pages 55-122
ISSN 2230-5734
https://doi.org/10.61618/DHXM4316

In partnership with

https://www.theicpem.org/

https://nasar.org/

www.isrid.net/

https://maritimesar.org/

https://www.landsar.org.nz/

Volume 8, Issue 2, December 2025
Pages 55-122
ISSN 2230-5734
https://doi.org/10.61618/DHXM4316

Contents

Volume 8, Issue 2, December 2025
Pages 55-122
ISSN 2230-5734
https://doi.org/10.61618/DHXM4316

JOURNAL of SEARCH+RESCUE

Editorial

https://doi.org/10.61618/UXQL2693

Welcome to Volume 8, Issue 2 of the Journal of Search and Rescue. This issue brings together work on the risks created by uncoordinated public drone use, new guidance for flight-testing SAR innovations, a proposed triage model for aquatic mass-casualty incidents, and a large-scale review of JESIP performance in UK incident command assessments, all offering advances in understanding of the challenges and emerging practice in SAR.

I want to use this editorial to mark the sad passing of our editorial board colleague and friend Toby Meredith. Many of you will have known Toby through his work on this journal, as an academic at the University of Portsmouth, his involvement in drone technology for disaster response and crisis mapping.

Toby started flying drones at Portsmouth about ten years ago and helped build the capability there from almost nothing. He worked in Dominica after Hurricane Maria, and later spent time in Senegal looking at seagrass with the Wingtra. Most of his projects sat somewhere between people, technology and hazards, using drones and data to make sense of risk and to help others plan for it. He also worked with the World Food Programme, travelling to Mozambique and later hosting WFP colleagues in Portsmouth. I was fortunate to spend two weeks in Africa with Toby on that project.

Some of you reading this will have been part of those trips or will remember planning them with him. The projects were serious scientific or response endeavours, but if you travelled or worked with Toby you will probably remember the laughter and the joy as much as the day job.

That is my strongest memory of Toby; obviously he was clever, capable and committed, but when I think back to Toby I hear the jokes. One of my favourite memories is in an evening in Maputo in Mozamonique, following two weeks of tough research, at a restaurant overlooking the Indian Ocean. The pressure was off, having completed the project data collection and we sat, shared a couple of bottles of wine some incredible seafood and chatted about our lives, our interests, our families, our ambitions and I remember feeling at that point how lucky I was to have him as a colleague and a friend

Inside the journal he was a key component, taking on important tasks behind the scenes, writing editorials and contributing articles to the journal. When he reviewed work, he did it with the aim of the publishing work, not providing barriers to publication. This is the central ethos of the Journal, and absolutely critical given our authors are often practitioners firstmost, and not from traditional academic backgrounds - but Toby embodied that ethos.

Some of Toby's work appears in this issue and some will follow in the next one. These papers were already moving through our normal process before he died. We decided not to attempt to turn them into tributes, but to leave them to stand as they are: pieces of work that tackle real issues in our field, and contribute to making the communities we serve safer. Future readers will just see good, useful, practical research.

Volume 8, Issue 2, December 2025
Pages 55-122
ISSN 2230-5734
https://doi.org/10.61618/DHXM4316

To all of you who knew and loved Toby, I know that an editorial cannot possibly carry the full weight of your loss, but I hope it helps in a small way to see how widely he was respected and how many people learned from him and enjoyed working with him.

I want to end with a simple thought. Toby's death is a hard reminder that our community is made up of real people with lives, families and limits. The work you all do in search and rescue, crisis response and disaster risk reduction is so, so valuable, but also demanding and often exhausting. So please, look after each other. Check in on your colleagues. Hug your loved ones. Embrace life, and laugh as much as you can. When you are doing what you do in these incredible places around the world, stop and think, soak in the environment and remember what a privilege it is to work in this field.

And as you continue your research, training and operations, stay safe. The work you are doing is important, and so are you.

Ian Greatbatch

Toby Meredith, Editor.
1982-2025

Considerations for flight testing of search and rescue (SAR) innovations

Christopher J Bennett, PhD[1], Jane Hodgkinson, PhD[2], Jim Nixon, PhD[2,3]

[1] Department of Aerospace Engineering, Embry-Riddle Aeronautical University, Prescott, AZ, USA
[2] Cranfield University, Cranfield, Bedfordshire, UK (at the time of collaboration)
[3] Health & Safety Executive, UK

Email bennec28@erau.edu

https://doi.org/10.61618/SITW4437

Abstract

Flight testing of new aircraft, equipment and systems is mandatory in the aviation industry to establish and verify the airworthiness of any new airborne product as determined by the appropriate authority, for example the Federal Aviation Administration (FAA) in the United States or the European Union Aviation Safety Agency (EASA) in Europe. Testing procedures for civilian and military category products are well documented and strictly regulated (see, for example the FAA flight test guide for certification of normal, utility, acrobatic, and commuter category aircraft (FAA, 2011) and often use traditional techniques. However, the are no published guidelines for flight testing new developments in the search and rescue (SAR) domain.

This article summarizes, from experience, recommendations for a successful SAR flight test campaign. This includes considerations for effective and efficient trial design, best practices for data capture across a range of environmental conditions, and methods to minimize sources of error. A detailed example of a search procedure with sequential steps is also provided. The principles discussed herein target quality data capture via flight test, both efficiently and safely, to improve and evolve the vital work that SAR operations conduct worldwide.

KEY WORDS Flight-Test design, innovations, review, case study

Introduction

Background

Airborne Search and Rescue (SAR) is a key element of emergency services worldwide. The objective of the SAR operation is to locate a casualty in the fastest possible time in order to complete a safe and successful rescue. When searching for casualties in challenging environments (rough seas or rugged terrain), SAR divisions typically deploy a helicopter, usually with a crew of four: Captain, First Officer (FO), rear search support crew, and a medically trained winch-man. Twin-engine, medium-lift helicopters are usually the aircraft of choice for such operations due to their ability to track low and slow over a search area, and also for their capability to hover and winch for the rescue phase. However, fixed wing aircraft are also commonly used to search much wider areas (or for larger targets) where faster, higher altitude flying is appropriate.

In the US, maritime SAR is typically conducted by the US Coast Guard who operate over 200 aircraft (both rotary-wing and fixed-wing) (U.S. Department of Defense, 2019). Inland SAR is typically coordinated by the US Air Force Rescue Coordination Center (AFRCC) and can leverage USAF, state, local or civilian resources, equipment and personnel (U.S. Department of Defense, 2023a). Deployment of resources, in either case, depends upon the severity/urgency of the distress call, availability of personnel and equipment, geographic location, terrain, and weather conditions. Between 2007 and 2017 the US Coast Guard conducted an average of 21,336 sorties per year, saving a total of 48,522 lives (Bureau of Transportation Statistics, 2017), and between 2013 and 2023 the AFRCC coordinated on average 750 missions per year, saving a total of 3430 lives (U.S. Department of Defense, 2023b).

The authors of this article previously completed a full-scale evaluation, including in-flight trials, of new conspicuity aids for casualties at sea for an EASA funded project (Bennett, Hodgkinson, & Nixon, 2019), and Bennett is currently involved in a SAR product design project which will soon be ready for ground and in-flight testing. This article has been motivated by the many discussions, issues that have arisen, and resolutions that have been devised in previous and on-going research campaigns.

Introduction of technology in SAR

SAR is a highly skilled, technology driven industry, continuously evolving to update equipment, processes, and procedures with the aim of improving the probability of detection, and therefore survival rates. Much of the equipment currently utilized by SAR operations has been developed specifically for this application. Furthermore, procedures such as search patterns, altitudes/speeds, and SAR crew resource management (CRM) have been developed to optimize the effectiveness of any given mission. However, there are a number of key technologies currently used by SAR operations which have filtered through other industries such as commercial aviation and military applications, for example with the introduction of the forward looking infrared (FLIR) system and night vision goggles (NVGs).

Clearly, the specific processes and procedures employed by any given SAR crew should be derived based on the equipment they have at their disposal. For example, if a new FLIR system has a wider swath than a previous model, the search pattern should be modified to take advantage of the improved capability and hence reduce the time taken to survey a given area. Therefore, as new SAR equipment is introduced, methods pertaining to the search procedure requires continual reassessment to fully implement the equipment effectively. This process obviously requires full-scale, in-flight testing and training. Ultimately, an evaluation as to whether a new innovation is worthwhile to implement from a cost and/or operational standpoint should be made. For example, a governing body is unlikely to invest hundreds of millions of dollars implementing a new technology on its fleet of aircraft if it is projected to save only a handful of lives over its operational lifespan. This is the unfortunate reality. However, if a new piece of equipment can be rolled out relatively cost effectively, and is projected to significantly improve the probability of detection and therefore the survival rate of a missing person, then the case for implementation is much stronger.

Aims and Organization of this article

This document details important considerations for designing and conducting a realistic and representative trial in the SAR domain, and also describes how many of the factors described in Figure 1 can be utilized to assess new technologies or working practices. The principles are applicable to research, governing bodies and private companies needing to assess the benefit, and airworthiness, of new innovations in a SAR context.

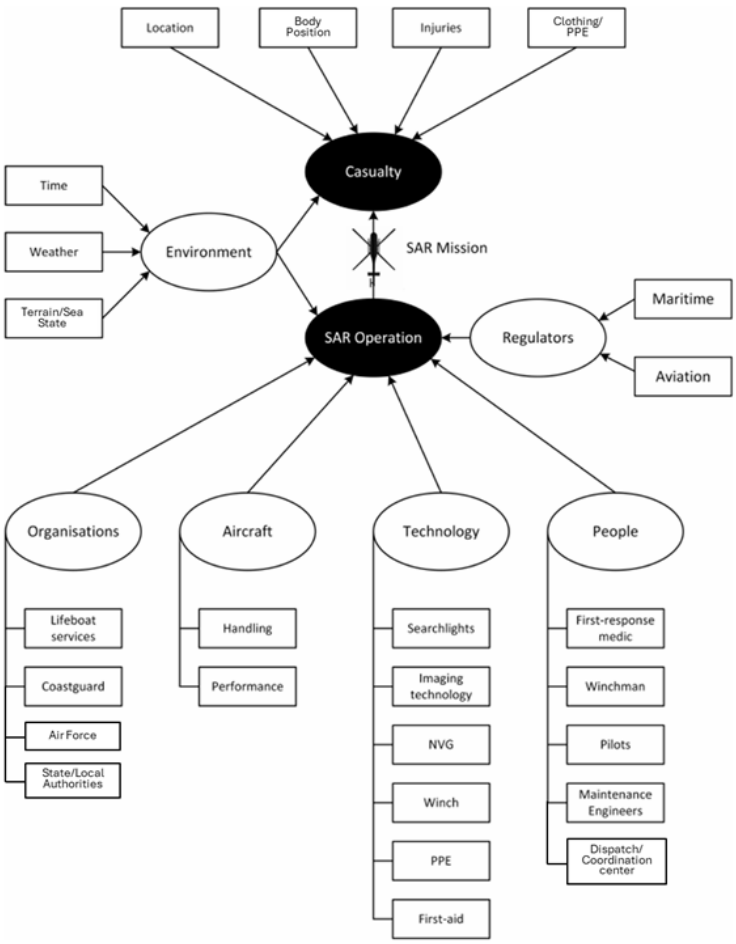

Figure 1: The SAR mission as a system of systems.

Aspects of the SAR mission considered in this article include:

- The performance and handling of the aircraft in support of the SAR mission.
- Imaging hardware and software including thermal imaging or the use of computer vision in SAR (Gotovac, Papic, & Marzusic, 2016), (Lygouras *et al.*, 2019).
- The color, intensity and type of aircraft search light.
- The rescue equipment and procedures used in a mission (for example the winch).
- The use of NVGs as an aid for the search and the optimal ways in which to use them.
- The effects of casualty clothing (including immersion suits, life jacket, and helmets on the mission.)

One key challenge when conducting research in this field is that any new design must be tested and validated in a real environment to ensure it is both safe to use, and also improves the SAR operation in a measurable and meaningful way. A comprehensive program of research is therefore required to support any hypotheses or claims made by equipment manufacturers. Typically, such research campaigns are conducted in the following sequence:

- Research into new SAR innovation: to understand the new capabilities, how it could be of benefit, and if it is possible to integrate the new equipment with the current operation.
- Research into current SAR operation: to understand current equipment capabilities, methods, procedures, and if the new SAR innovation is compatible with the current operation. This also establishes a benchmark for the new innovation to be tested against.
- Lab/ground testing and training: to test the new innovation against the benchmark in a controlled and safe environment. Measurements are taken and supporting data is gathered. The host SAR crew should also be trained on the new equipment in this phase, and feedback on its usage should be collected.
- Full-scale/flight-testing: to test the new innovation in a realistic environment. Measurements are taken and supporting data is gathered.
- Present findings: to justify the upgrade (or not) and outline the projected benefits and potential improvements to the SAR operation. Summarize the data and provide evidence such that the governing body can make an informed decision based on a projected improvement to cost ratio.

The protocol presented in this paper is derived from the authors prior experience, in particular, recommendations made to EASA (Bennett, Hodgkinson, & Nixon, 2019) currently under review/revision for an ongoing SAR product design project. Since no published guidelines, checklists, or instructions existed to aid with this specific type of project, the authors developed bespoke requirements. On reflection, this article represents a reference guide the authors hoped was available to consult at the start of this body of research. The lessons learned, trial design revisions, and protocol updates, as well as considerations of sources of error, have been distilled into the work presented in this article. The

findings are generalized such that they may be applied to assess any type of new equipment or operational method/procedure in a SAR context.
Here, an outline of principles and guidelines for the design and implementation SAR flight trials, are written to encompass a broad range of scenarios. The first topic of discussion is the design of the trials/tests. This will include a summary of best practices. Next, the myriad of factors which can harm the reliability and validity of the trials will be outlined, followed by a description of ways in which they may be resolved, or at least minimized. Finally, a step-by-step procedure that could be applied to any SAR based trial is provided as an example.

Approaches to SAR trials – A review of previous studies

Different approaches have been taken by researchers in various fields to the problem of assessing the efficacy of new innovations in SAR. These primarily focus on maritime SAR. The type of testing, and the design of the trials, is highly dependent on the type of product in question, for example, aircraft hardware, new imaging technology, drift models, algorithms for defining and programming the aircraft's search pattern, and unmanned autonomous vehicles (UAVs).

For a complex application such as SAR target detection, a full-scale test is required to validate the operation of the technical systems and the interaction with human operators and observers. However, to do this thoroughly, a sufficient number of realistic SAR scenarios would be required to enable statistically significant comparisons between different options to account for system reproducibility across a number of variables: lighting conditions, weather and visibility, sea states, orientation of target object, plus human factors such the level of alertness of operators and their ability to discriminate the target from clutter in the field of view. Some of these may be chosen but still have variability: the weather and lighting might be broadly similar but can still vary between one test and the next. There are additional purely stochastic factors such as whether the crew were looking in a particular direction at one time or another, or whether a mission's success was due to happenstance rather than by systematic design and execution. Clearly, the number of missions that would be required to control, or average, over such factors is unfeasibly large if conducted under realistic conditions, at scale. It is also difficult to run many repeated scenarios in the same location without the same human operators drawing on knowledge of previous runs: where to look and what the target looked like under the prevailing condition. Despite these limitations, well-considered trials and comparisons can be usefully made. Different types of testing provide complementary information, as discussed below.

Specification testing

For certain technical subsystems, it is possible to define a set of operating requirements and specifications that any given new innovation must fulfill. Hence, the new technology may be tested quantitatively against the corresponding requirement. This also includes testing of models and algorithms used in planning missions. The key here is that a single part of the whole SAR system is evaluated in situations that have reduced complexity. Subsystems that have been tested in this manner include control systems for UAVs (Almeshal & Alenezi, 2018), where it is possible to identify the ideal flight path and measure deviations from that ideal.

The advantage of such testing is that it may be readily controlled and replicated, which is essential to the development of technical solutions particularly in the early stages. The limitation is that it can only be applied to subsystems that are separable and well understood. In (Bennett, Hodgkinson, & Nixon, 2019), for example, performance of retro-reflective materials that may be applied to crew survival suits were characterized optically in the laboratory. Although performance standards exist for these materials (International Maritime Organization, 1989), (American Society for Testing and Materials, 2013), they relate to one type of lamp and one standardized spectral response of the detector only. Therefore, additional tests looking purely at optical performance, in ground-based trials of detectability, were required.

Detectability

Measurements of target detectability can be made during trials on sea or land using the same or similar detection equipment to that deployed on SAR aircraft. For this type of trial, the location of the target should be known, but the assessment is of the strength of the detection signature under different circumstances. `Strength of the signal', here, may refer to the distance at which the target is observable by the SAR crew, the contrast of the target within its environment, or both. See, for example, the tests conducted in (Burciu, Abramowicz-Gerigk, Przybyl, Plebankiewicz, & Januszko, 2020) aimed at enhancing heat signatures of life rafts, the tests of a novel detection system based on sensor fusion conducted in (Burciu, Abramowicz-Gerigk, Przybyl, Plebankiewicz, & Januszko, 2020), and the investigations into immersion suit conspicuity performed in (Hodgkinson, Nixon, Bennett, & Tatum, 2020) and (Nixon, Hodgkinson, & Bennett, 2020). In all cases, a stronger signal should lead to a higher probability of detection within a given time.

Drift models

Drift is the term used to describe the motion of floating objects due to wind and sea conditions. By knowing the location at which the target entered the water, and the weather and sea conditions, mathematical models can be applied to predict the change in position of the target over time. Numerous field trials have been performed to assess drift models that aim to identify the likely location of targets and thereby narrow the search area for the SAR crew. Trials are typically conducted with standard buoys as well as real targets in different seas, to validate the reliability of the mathematical models and meteorological inputs to handle the local conditions, for example the tests conducted in (Cho et al., 2014). Trajectories are compared over large datasets to analyze accuracy and reproducibility, in a similar way to weather forecasts, for example. In (Coppini et al., 2016), known initial and final recovery locations of drifting objects such as ships, boats, and people lost overboard were used to assess common drift models. However, good quality, reliable data is often too sparse to permit statistically significant evaluations, as discussed in (Breivik & Allen, 2008).

Realistic field exercises

Exercises should be designed to simulate a real SAR event, in order to test the function of a whole system or at least the part of the system involving SAR operators. In the context of maritime SAR, researchers often use a weighted raft or SAR training dummy/manikin as a search target rather than human participants, for reasons of safety. This cannot account for differences in behavior of live casualties, for example waving to an approaching aircraft to attract attention. Furthermore, resources to conduct trials with `live' SAR crews are limited, often requiring volunteers and/or integration with training exercises. Due to the practical and operational limitations, the amount of raw data which can be collected in any given trial is severely limited. Hence, there are very few published trials which can be considered to test all or a large part of the SAR system is a realistic way.

Despite this, there have been several key studies involving full scale SAR trials which have led to meaningful conclusions and, as a result, operational changes. For example, in (Donderi, 1994) a total of 116 Coast Guard crew members on-board two ships were tasked with identifying 23 tethered life rafts using the unaided eye, binoculars and NVGs for a period of one hour. From such a large-scale realistic test within a realistic environment, it was possible to capture a wealth of statistically significant data. Other full-scale trials have involved investigating UAVs for SAR in a post-tsunami situation (Ferreira et al., 2018), life jacket tracking devices (Lilja et al., 2013), (Miano et al., 2019), and helicopter visualization systems (Miller, Kelly, & Ehler, 1999). It is interesting to note the differing approaches to `measuring' the success of each trial. The performance metric needs to be carefully considered so that the results can be ultimately related to the probability of detection. This will be discussed in more detail later in this document.

It is therefore evident there is a lack of published research using a full-scale SAR operation, and hence forms the motivation for this article.

SAR trial design principles

Trial goals

Any given trial should be aimed at improving one of the overall goals of the SAR operation:

- **Prioritize:** using local knowledge and computer software (to calculate drift, for example), to identify the search area.
- **Search - Detect:** improve the probability of detection of objects.
- **Search - Identify:** recognize and categorize whether the detected object is a target of interest.
- **Rescue:** improve the efficiency of the rescue of a casualty.
- **Safety:** improve safety for the crew and/or the casualty.

The goal should be well defined, specific, and quantifiable to determine if the goal was achieved.

Trial management

The research team are responsible for planning the trials and managing the project. Clearly, they should work in close collaboration with the other parties involved to ensure that the objectives are:

- achievable given the specific equipment available,
- reasonable given the integration with a live SAR operation,
- safe for all parties involved.

Therefore, the research team personnel should be selected to include experience and expertise in the subject area. This may include, human factors/CRM, flight operations, flight test, planning/scheduling, communications, flight mechanics, sensor technology, software, optics, etc. This will gain the project credibility and allow the best use of limited flight time available. The synergy of the research team, whether large or small, is key. They should have a broad range of skills, expertise and experience with regards to the scope of the project. Interaction and effective communication between the research team and the SAR crew is also paramount to disseminate goals, ideas and findings at each step of the project. The research team are also responsible for detailed and professional mission briefings for the SAR crew.

Participants

It is important for the research tasks to be performed by trained individuals, familiar with the current methods and procedures. It is important that participants are recruited from a genuine population of working SAR crew, to provide a realistic test. Crew should also not be required to operate in unfamiliar roles. Again, these considerations are aimed at obtaining as realistic conditions as possible for the tests, and maintaining scientific validity of the processes and procedures. This will ensure that the best quality data is captured from the trial.

Performance metric

The chosen performance metric should allow results to be measured scientifically, and lead to quantitative conclusions. It is often difficult to isolate specific factors from a real-world event and hence the trials should be designed in a way to eliminate variability as much as possible to capture consistent data.

There are typically two types of performance metric which can be measured in this context: a simple binary measure of success or failure (for example, target identified or not), or a continuous numerical value (for example, distance or time). The binary measure requires a significant number of tests to be completed to satisfactorily assess the difference in performance between the benchmark verses the new innovation. In contrast, measuring a continuous variable (for example, the distance to the target at which the object is first detected) provides more granular data, and hence differences between variables can be derived from a reduced number of tests. This metric can be directly related to the time required to search a specified area (since the longer the distance at which the target may be detected, the wider the search swath of the aircraft) and thereby the probability of survival of casualties [12]. In the search phase, the success of a trial may be measured by the time taken to locate the casualty, and is a parameter considered in trial planning software.

However, for a limited number of flight trials the time taken to detect the target may be strongly influenced by luck: the target might happen to be located early or late in the chosen flight trajectory. The problem is avoided if the distance at which the target is first detected is measured, rather than the time taken to identify. This can be achieved conveniently by `dropping' coordinates using the aircraft's GPS system, for example. The distance and time parameters are easily linked; the greater the detection distance, the wider the swath of the aircraft sensors and the shorter the overall flying time needed to cover a given area (Burciu, Abramowicz-Gerigk, Przybyl, Plebankiewicz, & Januszko, 2020).

Data gathering

Having decided upon a suitable performance metric which will enable direct comparisons with the control condition, the question becomes how to accurately capture the data with the resources available. Typically, in-flight research uses sensors and instrumentation to record information in real time. This alleviates workload on the pilot/crew on-board the aircraft. Clearly, it would not be appropriate to expect a pilot to manually record data while in control of an aircraft for safety reasons. Furthermore, the operation should be conducted in as close to realistic conditions as possible, i.e. without added pressures, tasks, or distraction. Regarding the aircraft itself, it may be possible to add additional equipment for the purpose of the flight test, or modify the aircraft in some way to help capture the necessary data. However, depending on the type and intrusiveness of the modification, airworthiness assessment and certification may be required. Hence, the design of the flight test should take advantage of any existing equipment/sensors wherever possible.

An example of an efficient strategy for data capture during a SAR trial is to record the FLIR video of the entire search with audio from the aircrew intercom (IC) overlaid. Throughout the US and Europe SAR aircraft are typically fitted with the FLIR system and the video/audio recording feature should be available as standard. The FLIR video recording contains live flight information including ground speed, flight path and GPS location. In combination with the IC, GPS coordinates at the point when the target was first identified, and the GPS coordinates of the target itself, could be recovered following the trials via analysis of the video. From experience, this is a convenient method for retrospectively calculating the identification distances following the trials and was minimally obtrusive to the crew, allowing them to operate in a natural manner. It must be noted that FLIR video/audio recordings may be proprietary and/or protected against release to the public. This should be respected by the research team when presenting the findings.

Due to the significant variability (planned or unplanned) that could occur during a SAR trial, particularly regarding the environment in which the search takes place, there is also important supplementary data that should be recorded (see the lists below). Checklists are recommended to ensure integrity of the data in a complex trial where details could easily be forgotten. This additional data should be captured either at base before take-off, or by the non-flying pilot during the test. Again, considerations must be made to ensure that the requests for data logging by the crew are reasonable, and do not interfere with the basic operation during flight.

The following details should be recorded (see (Bennett, Hodgkinson, & Nixon, 2019) for an example test card) by a member of the aircrew **prior to take-off**:

- Date and time (if retrospective information regarding weather/sea conditions is required)
- Aircraft Type and Registration (for information on performance and equipment)
- Crew (for analysis of experience levels - hours, aircraft types, training, military/commercial background etc.)

The following details should be recorded (see (Bennett, Hodgkinson, & Nixon, 2019) for an example test card) by the non-flying pilot **prior to the start of each search pass**:

- Trial Number and configuration (for correlation with the data collected)
- Conditions (for analysis of effects on the success of a trial) including: air temperature, ambient pressure, visibility, sea temperature, wind speed and direction, weather, sea state, light meter reading (this can be communicated from a ground-based member of the support team).

Raw data should also be supplemented with comments and feedback from the aircrew in a debrief back at base. It is beneficial for the researchers to attend the flight debrief (where possible) and interview the crew to document their experience immediately following the trial. A combination of qualitative and quantitative data allows the researchers to conclude whether the innovation being tested offers a benefit in comparison to the current standard.

It is also important to prioritize the data which is sought so that critical data is collected first, followed by any supporting/repeated data if the resources and opportunities allow. It is a case of managing expectations and being realistic with what can be achieved given the resources available. It is common for trials to go awry when integrating research with a live operation. For example during the testing phase for (Bennett, Hodgkinson, & Nixon, 2019), while integrating a research trial with the SAR crew's scheduled training exercise, the crew received a real distress call and had to abandon the test mid-trial. This led to the loss of usable comparative data with respect to the control condition which had already been completed, and also the loss of the training manakin which was fortunately recovered the following day by the local lifeboat. As a further example, in a later trial for (Bennett, Hodgkinson, & Nixon, 2019), rough sea conditions prohibited the targets from being deployed in the search area originally planned for the test. This meant that the SAR crew, already in the air and ready to search, were briefed incorrectly. The trail, and subsequent data, was recovered by repositioning the targets into a more sheltered area of coastline. The SAR crew were updated via radio of the new search procedure and were able to successfully complete the trial in full, albeit under conditions which could be considered marginally less representative of a real SAR event.

To ensure maximum validity of the data, the SAR crew must employ methods and equipment as defined by the most up-to-date International Aeronautical and Maritime Search and Rescue (IAMSAR) guidelines (IMO Publishing, 2019).

Controls

It is important to test any new innovation directly against a control to quantify any potential improvements and ultimately make recommendations. The most valid approach in most cases is to use the current industry standard as a baseline comparison. It should be tested at the same time and in the same conditions in order to provide a meaningful conclusion based on the performance metric (discussed above). For example, in a previous project by the authors (Bennett, Hodgkinson, & Nixon, 2019) regarding survival suit conspicuity, a standard unmodified immersion suit was chosen as the control measure.

Based on a review of the literature in the field of SAR flight testing, it appears that many published studies have not explicitly evaluated a direct comparison of the control, rather concentrating on validating the efficacy of the new innovation in isolation. As a result, it is difficult to make solid evidential recommendations. However in some cases, controls are not possible, for example when testing location beacons Lilja et al., 2013).

Blinding

Another consideration when conducting multiple trials of the same type, with the same participants, and in the same location is `blinding'. In this context, blinding concerns eliminating the influence of previous trials on the current one, for example, the participants gaining experience of the experimental set-up. This is somewhat unavoidable when testing against a control as discussed above. Hence, it is important to design the trails in such a way to minimize this effect. Furthermore, `double blinding', which concerns eliminating the influence of the scientists designing the trials on the participants of the trial, should also be considered. Clearly, there is an implication of blinding on the repeatability and reproducibility of any given trial, which is discussed further in the next section.

Again drawing upon experience, the key here is to design the tests in such that the SAR operation gains minimal transferable knowledge and experience from one test to the next. For the authors previous project regarding immersion suit conspicuity (Bennett, Hodgkinson, & Nixon, 2019), the trials were designed such that the aircraft's flight path would always traverse directly over the search area, and that the target would always be within the scanning arc of the pilots and FLIR system. Therefore, the metric for success (the distance at which the target was first identified) was intentionally not dependent on how the aircraft was flown. It was specified that, having completed a run, the aircraft was to continue past the target to a landmark sufficiently far from the search area such that it would be impossible to identify the target, before turning around to make the next pass. There was, however, some variation in the location of the target within the search area due to tidal drift between passes, and also due to random nature of how the support boat deployed new targets within the search area. It was found that in most of the trials conducted, the visibility in the prevailing conditions was so limited that the crew were unable to make much use of knowledge gained on prior runs. The results show that the crew's performance on repeated runs was remarkably similar and did not noticeably improve as more trials were conducted. This is the goal in maximizing blinding.

Despite these successes in (Bennett, Hodgkinson, & Nixon, 2019), double blinding was difficult to achieve fully, as the research team were responsible for placing the target object on the sea surface and therefore knew its location. However, there are restrictions forbidding untrained personnel from being on-board the SAR aircraft, particularly when flying over water. Only SAR crew were allowed to be present and therefore scientists running the trial were explicitly not present when the helicopter was in the air. Therefore the research team are unable to directly or indirectly influence the crew's actions during the search phase. Radio communication between the team on the surface and the air crew was limited to events that took place after target detection occurred, or (in some cases) the search was abandoned.

Blinding in trial design is not frequently discussed, though it may have been implemented; for example in (Donderi, 1994) the measures taken to ensure that participants were unaware of object locations is not discussed. However, it is also important to ensure that a trial remains tractable and easy to conduct within a reasonable time window, therefore the target location must be somewhat constrained.

Repeatability and reproducibility

In a realistic field trial, it is difficult to assess the repeatability and reproducibility of a measurement while also ensuring blinding of participants and a realistic participant population. For example, it is typically not practical or efficient to change the search area, or to significantly change the position of the target within the search area, for each run due to the limitations of resources and time constraints. However, efforts should be made to blind the SAR crew as best possible from run to run as described above. Furthermore, there is typically not enough potential participants within the population to send out a new crew for each run.

Repeatability and reproducibility is also difficult to achieve regarding environmental factors such as weather, light conditions, visibility, sea state, sun position with respect to the direction of flight, etc. The key is to minimize the time required for each run so that the change in prevailing conditions is minimized from run to run. In (Bennett, Hodgkinson, & Nixon, 2019), the average trial time was 6-8 minutes and so any change in conditions could be considered negligible. Therefore, the dependency on the conditions can be removed from the comparison of the data.

It is recommended that if the trails are likely to take greater than 10-15 minutes, weather and sea conditions should recorded by the SAR crew prior to each trial. This can be completed based on the weather radar at base, and the Automatic Terminal Information Service (ATIS) at the station airport. This should include sea state and tidal phase, lighting (cloud cover, position of sun or moon, etc.) and prevalence of rain, cloud or mist. For longer duration trials, any significant change in the conditions listed above during the test should also be noted, or otherwise considered constant.

Best practices for maximizing repeatability and reproducibility as employed for the maritime SAR trials in (Bennett, Hodgkinson, & Nixon, 2019) include choosing the trial time to be within a 2 hour window of slack water (high or low tide) to minimize the change in sea state and drift, and when the weather conditions were not predicted to significantly change. Differences in lighting conditions (ie. position of the sun/moon, and glare from nearby towns/cities) and the physical orientation of the target in the water should be mitigated by performing two runs in opposite directions. To assess reproducibility over a range of environmental variables, trials should be conducted in different locations and at different times of the year, in different sea states, wherever possible.

Sources of inconsistencies, interference and bias

There are a number of factors which could introduce inconsistencies, interference and systematic bias to the results of full-scale SAR simulations, confounding the validity of the trial. To allow comparison of results from multiple flight trials, it is important that these effects were mitigated or minimized as best possible. Below is a summary of typical sources of error, type, and best practices for mitigation/minimization. This is by no means a comprehensive list. The highly specific nature of any test will inevitably lead to a new and complex series of potential sources of inconsistencies, interference and/or bias. The subsections below summarizes the most common and potentially adverse examples when conducting tests specifically in the SAR flight test domain.

Visual scanning range of the SAR crew

This is considered a potential inconsistency. This can be mitigated/minimized, for example, by maintaining the following scanning ranges constant throughout the testing: FO 10-1 o'clock, Captain 11-2 o'clock, rear crew 1-5 o'clock, FLIR operator 9-3 o'clock.

Position and scanning motion of the searchlight

This is considered a potential inconsistency. This can be mitigated/minimized, for example, by maintaining the searchlight fixed at 12 o'clock, fully de-focused.

Multiple targets

This is considered a potential source of interference. This can be mitigated/minimized by testing the new innovation and current standard separately, in the same environment, in the same conditions.

Search crew becoming familiar with the test procedure

This is considered a potential source of bias. This can be minimized by employing different search crews and a range of experience levels wherever practically possible.

Search crew becoming familiar with the search area

This is considered a potential source of bias. This can be minimized by scheduling a number of search trials in different locations wherever practically possible.

Randomness of success in identifying targets

This is considered a potential inconsistency. This can be minimized by conducting multiple tests of the same configuration.

Pretesting, efficient use of resources, and financial constraints

It is clear that from the considerations discussed above that there is a delicate balance to be found between scientific rigor and the practicality of integrating trials with a live SAR operation. Any such tests obviously puts a strain on resources, relies upon cooperation (and often volunteers), and can be very costly. Utilizing crewed helicopters as the SAR platform costs approximately $10,000 per hour (excluding crew time). From a research perspective, it is key to realize that the day-to-day operation will always be prioritized at any base, potentially limiting the time window during which tests can be conducted.

Flight testing should therefore be the absolute final stage for any study of this kind. By its very nature, flight testing is expensive and comes with risks. Time spent in the air gathering data must be fully utilized, and the trials should be meticulously designed in such a way that the maximum amount of scientific data can be extracted from a minimal amount of flight hours. Time is critical and therefore it is not realistic to expect a large number of repetitions. Only once all constraints have been considered can a test plan be formulated.

Therefore, prior to any flight test planning, all necessary laboratory and ground testing should be completed to limit the number of variables that need to be tested at full scale. This should include, in particular, compatibility checks of all equipment with the proposed innovation. The goal is to clearly justify and constrain the flight tests by using the data from the lab/ground testing as evidence. This will minimize the number of discrete tests, and repetitions of tests, to be conducted at full scale.

The probability of success of a trial should also be considered. For example, is it likely that the data/evidence gathered from any given test will substantiate the claims/goals of the project? If the answer is yes, then the full-scale trials can be justified in terms of the resources required. If no, then the flight tests should be redesigned to improve the probability of success within the given constraints.

To alleviate (or absorb) some of the costs associated with flight testing, there may be operational opportunities to exploit. For example, SAR crews regularly undertake mandatory training exercises and so the research team may be able to integrate the field trials with these objectives. This was the case for the tests conducted in (Donderi, 1994) and (Bennett, Hodgkinson, & Nixon, 2019).

Trial location

The location of any trial should be chosen based on a number of factors: safety, representative conditions, financial, convenience, availability of equipment and personnel.
In the authors previous project (Bennett, Hodgkinson, & Nixon, 2019), ground and flight trials were conducted at 2 different SAR bases in the UK. It is important to recognize that SAR operations will always take precedence over any research related trials and therefore the research team should be prepared to spend long periods waiting for access to operational aircraft or crew members. A private research office, if available, can be used as a base, allowing the team to work effectively and also to minimize disruption to the SAR operation. For safety reasons, SAR officers are required peace and quiet during rest periods and equipment maintenance takes place in a no distraction zone. Ultimately, it is important that the research team respects normal SAR working practices.

Locations for the flight tests themselves should be chosen to be safe and un-congested in terms of both airspace and shipping. The search area should be large enough so that the exact location of the target is not trivially obvious to the SAR crew, but should be confined to the defined scanning ranges of the crew when following the predetermined flight track. Any landmarks should also be sufficiently far from the search area so that the target was not trivially visible from the turning points.

Local authorities should be informed such that the search trial is not mistaken for a genuine SAR mission. If the flight trial is to be conducted near a shore, it is a good practice to utilize a member of the research team to reassure and inform bystanders. These events are likely to attract attention and potentially cause distress to the public if there is a lack of information. Also, there is potential for outside personnel to interfere with the test in progress. For example in (Bennett, Hodgkinson, & Nixon, 2019), during one of the night-time trials, a member of the public parked their car on the promenade and directed the main headlight beam toward the beach from where the lifeboat crew were supporting the trial. Although this person was clearly trying to assist with a potential genuine rescue attempt, this type of occurrence would have been detrimental to the validity of the trial data, particularly in the case of evaluating the conspicuity of a casualty at sea from the air. In this case, the ground-based member of the research team was able to reassure the individual, with the result that they turned off their headlights.

Trial briefing

A briefing between the parties (for example, SAR crew, support boat crew, and the researchers/observers) should take place before each trial. A date and time for the trial should be decided, taking into consideration the requirement for a range of conditions as described above.

It is important for all involved to remain flexible and adaptable in the case that the trial does not go as planned. Appropriate contingency planning should ensure that time in the air is not wasted. A plan B should be formulated and briefed before each flight trial to give the commander the ability to change the objective of the trial mid-flight if it was deemed not safe to continue with plan A. This decision lies ultimately with the aircrew in this situation, as they are most at risk, and should be respected. A working radio frequency should be established so that the SAR crew can communicate with the research team (and support boat, for example, if applicable). Clear and efficient radio communication is key for a successful trial.

As an example of adaptability while on trial, during the second series of flight tests in Port Talbot, Wales in (Bennett, Hodgkinson, & Nixon, 2019), the plan initially was to conduct a search several miles out to sea, near to a tethered buoy which was used to signpost a shipping lane. Due to severe sea conditions and bad weather, deemed unsafe for the support boat, the search area was changed to a sheltered bay. As a result, the SAR crew had to change the aircraft's flight path to track parallel with the coast rather than in and out to sea. This decision allowed the boat crew to operate more safely, while maintaining continuity and validity of the trial.

Trial procedure and execution

While each SAR trial is unique depending on the overall objective and the data sought, the following example flight test methodology highlights the level of detail required. It is important to link the procedures of all parties involved with focus on communication. For example, the search helicopter must be aware of the support boat status at all times and maintain live communication via radio. The research team on the ground should also monitor these radio communications. This is necessary to ensure safety and efficient completion of the flight test sequence.

Effective and efficient teamwork, as well as clear and decisive communication is essential. Figure 2 shows a helicopter and support boat working together at sea. Conducting a flight trial is not without risk, so it should be performed exactly as planned, where every person involved understands their role and is confident in performing under pressure. Clearly the aircrew are experienced in operating under these conditions, however the ground support and research team may not be. The key message is that meticulous planning can ultimately achieve the research goals while saving time, effort, money, and also minimizing risk for all involved.

Figure 2: Effective teamwork between the SAR helicopter and the support boat during a flight trial, Bristol Channel, UK. Note that the image has been edited to disidentify the aircraft and the winchman sitting in the bow of the support boat.

Figure 3 shows a schematic for reference with Table 1 which sequentially lists the procedure for one of the flight tests conducted in (Bennett, Hodgkinson, & Nixon, 2019), as an example. This procedure applies specifically to the testing of new equipment in flight. However, if minor changes are made to the methodology, this strategy may also be used to test more conceptual SAR processes such as pilot scanning ranges, aircraft search patterns, and pilot training techniques.

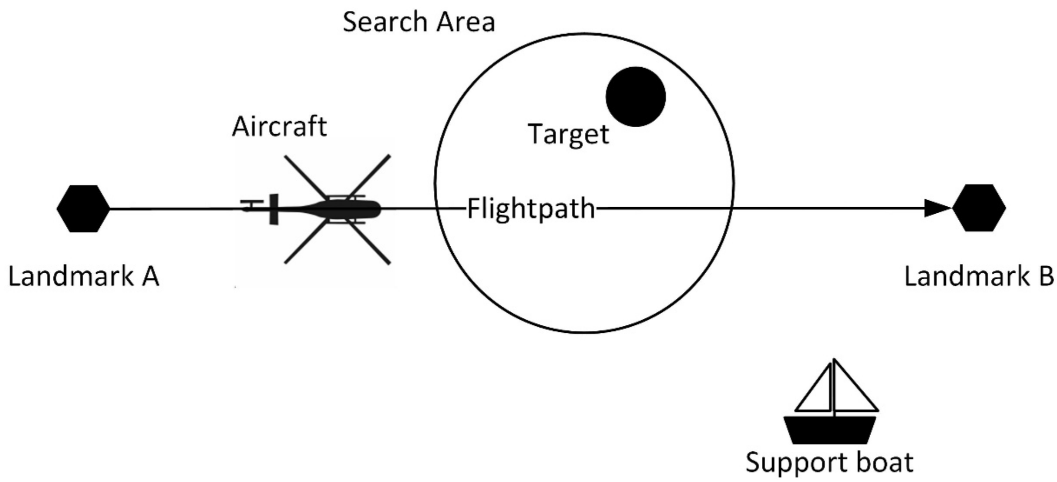

Figure 3: Schematic of search for the example given in Table 1.

1	Boat launches and transits to search area.
2	Boat deploys first target with PLB attached.
3	Boat transits to stand-off area and makes a radio call to the aircraft `ready for search'.
4	Pilot completes the first portion of the test card.
5	Aircraft takes off and transits to landmark A.
6	FLIR operator starts the video recording.
7	Non-flying pilot completes the test card for trial 1.
8	Aircraft makes first pass from landmark A to landmark B searching for the target. On identifying, the crew member who spots says `spotted from cockpit/on FLIR'.
9	Aircraft transits to directly above the target and the pilots says `on top'.
10	Aircraft continues to landmark B and turns around.
11	Aircraft makes second pass from landmark B to landmark A searching for the target. On identifying, the crew member who spots says `spotted from cockpit/on FLIR'.
12	Aircraft transits to directly above the target and the pilots says `on top'.
13	Aircraft makes radio call to support boat `second pass complete'.
14	If the target was not found, the SAR crew may use the PLB to locate and aid the recovery by boat.
15	Support boat retrieves the first target and deploys the second.
16	Aircraft continues to landmark A and turns around.
17	Support boat transits to stand-off area and makes radio call to aircraft `ready for search'.
18	Non-flying pilot completes the test card for trial 2.
19	Aircraft repeats the two search passes.
20	The trial procedure repeats as necessary.
21	On completion of the final pass the FLIR operator stops the recording.
22	Support boat retrieves the target following the final pass and returns to dock.
23	The aircraft transits back to base, FLIR video is downloaded, and the crew debrief as normal recording any further details pertinent to the trial.
24	Research team performs the necessary analysis of the FLIR video recording.

Table 1: Example search procedure.

Data analysis, results and recommendations

Assuming that the flight tests are successful and sufficient data has been collected, the research team can proceed to perform the required analysis. The first step is to quantify the results in context of the control condition. This includes assessing the validity of the data across the expected full range of conditions in a real SAR scenario. Both quantitative results (based on the chosen metric) and qualitative/anecdotal results (provided by the SAR crew in debriefs) should be presented.

Based on the analysis of the results, realistic and practical recommendations should be made. For example, the financial implications of introducing a new standard, in terms of the cost of both the new equipment itself and any additional crew and maintenance training (initial and recurrence) required, should be considered. Generally, it is not the role of a research team to recommend that SAR operations change, but to provide an evidence base that can be used to inform decisions that may, for example, balance implementation costs and quality of response.

As an example, in the authors previous project (Bennett, Hodgkinson, & Nixon, 2019) focusing on crew immersion suit conspicuity, it was recommended that additional retro-reflective tape, positioned strategically on the survival suit, should be employed. The recommendation represents excellent cost effectiveness: the material itself is relatively inexpensive, the suits do not require radical redesign, no additional maintenance is required, and it was demonstrated that the modifications are compatible with existing SAR equipment on-board the aircraft. It was also shown through additional ground testing that the modified suit design did not lead to unwanted reflections, distraction, or discomfort within the cockpit. Therefore, the project and overall findings were considered successful in completing its objective by EASA, with the recommendation that regulators, immersion suit manufacturers, and SAR should consider the proposed design modifications to aid conspicuity at sea.

Conclusions and final remarks

In light of no published guidelines specifically for SAR flight-testing, and also with the noted lack of published articles related to full-scale SAR testing, this article presents the key considerations for designing and conducting a successful campaign, drawing upon prior experiences and lessons learned.

In the first half of this article, a review of the literature is provided to highlight the differing approaches to SAR testing. While a range of platforms, targets, locations and methods have previously been studied, published articles in this field focus mainly on a single component of the SAR operation, and often not at full scale or representative conditions. In addition, no previous studies consider a full-scale SAR flight trial using a helicopter (the most widely used method) probably due to the expense and availability.

In the second half, each SAR trial design principle is discussed to highlight the constraints, as well as methods to improve the validity of the trials, and ultimately the data required to solidify recommendations. For each section, examples are provided based on published studies or prior experience of the authors, primarily in the context of maritime SAR. However, these basic principles are relevant to any SAR flight test campaign, and it is the intention of the authors that this article can be used as a reference guide for future research projects in this field. It is hoped that the article can be used to ensure that the very best possible data can be captured in the most efficient and safe way. This article has been motivated by the many discussions, issues that have arisen, and resolutions that have been devised in previous research campaigns. The authors hope that the findings are of interest to the flight test community and most importantly can be used to improve and evolve the vital work that SAR operations conduct across the world.

References

Federal Aviation Administration (FAA) (2011). Flight Test Guide for Certification of Part 23 Airplanes. US Department of Transport, No. 23-8C.

Bennett, C. Hodgkinson, J. Nixon, J. (2019). Crew Immersion Suits Conspicuity. European Union Aviation Safety Agency (EASA), No. EASA_REP_RESEA_2017_2.

U.S. Department of Defense (2019). Factsheet. Online: https://www.uscg.mil/datasheet/display/Article/1547982/aircraft/, accessed June 26 2025.

U.S. Department of Defense (2023a). Air Force Rescue Coordination Center – Mission. Online: https://www.1af.acc.af.mil/, accessed June 26 2025.

Bureau of Transportation Statistics (2017). U.S. Coast Guard Search and Rescue Statistics. Online: https://www.bts.gov/content/us-coast-guard-search-and-rescue-statistics-fiscal-year, accessed June 26 2025.

U.S. Department of Defense (2023b). AFRCC 2023 Annual Report. Online: https://www.1af.acc.af.mil/Units/AFRCC/Annual-Reports/, accessed June 26 2025.

Gotovac, S. Papic, V. Marzusic, Z. (2016). Analysis of saliency object detection algorithms for search and rescue operations. 24th International Conference on Software, Telecommunications and Computer Networks (SoftCOM), No. 16522518.

[8] Lygouras, E. Santavas, N. Taitzoglou, A. Tarchanidis, K. Mitropoulos, A. Gasteratos, A. (2019). Unsupervised Human Detection with an Embedded Vision System on a Fully Autonomous UAV for Search and Rescue Operations. Sensors, Vol. 19(16), No. 3542.

Almeshal, A. Alenezi, M (2018). A Vision-Based Neural Network Controller for the Autonomous Landing of a Quadrotor on Moving Targets. Robotics, Vol. 7, No. 71.

International Maritime Organization (1989). Use and fitting of retro-reflective materials on life-saving appliances. IMO, Resolution A.658(16).

American Society for Testing and Materials (2013). Standard Practice for Measuring Photometric Characteristics of Retroreflectors. ISTM International, Active Standard ASTM E809.

Burciu, Z. Abramowicz-Gerigk, T. Przybyl, W. Plebankiewicz, I. Januszko, A. (2020). The impact of the improved search object detection on the SAR action success probability in maritime transport. Sensors (MDPI), Vol. 20, No. 3962.

Hodgkinson, J. Nixon, J. Bennett, C. Tatum, R. (2020). Field investigation of retroreflective materials for enhanced target detection in maritime search and rescue. Proceedings - Optical Sensing and Detection VI, Vol. 11354, No. 61.

Nixon, J. Hodgkinson, J. Bennett, C. (2020). Modified immersion suits for helicopter aircrew: Evidence for improved conspicuity from sea trials. Safety Science, Vol. 130, No. 104903.

Cho, K. Li, Y. Wang, H. Park, K. Choi, J. Shin, K. Kwon, J. (2014). Development and validation of an operational search and rescue modeling system for the Yellow Sea and the East and South China Seas. Journal of Atmospheric and Oceanic Technology, Vol. 31, Pg. 197-214.

Coppini, G. Jansen, E. Turrisi, G. Creti, S. Shchekinova, E. Pinardi, N. Lecci, R. Carluccio, I. Kumkar, Y. D'Anca, A. Mannarini, G. Martinelli, S. Marra, P. Capodiferro T. Gismondi, T. (2016). A new search-and-rescue service in the Mediterranean Sea: a demonstration of the operational capability and an evaluation of its performance using real case scenarios. Natural Hazards and Earth System Sciences (NHESS), Vol. 16, Pg. 2713–2727.

Breivik, O. Allen, A. A. (2008). An operational search and rescue model for the Norwegian Sea and the North Sea. Journal of Marine Systems, Vol. 69, Pg. 99–113.

Donderi, D. (1994). Visual acuity, color vision and visual search performance at sea. Human Factors, Vol. 36, No. 1, Pg. 129–144.

Ferreira, F. Ferri, G. Petillot, Y. Liu, X. Franco, M. Matteucci, M. Pérez-Grau, F. Winfield, A. (2018). Scoring robotic competitions: balancing judging promptness and meaningful performance evaluation. IEEE International Conference on Autonomous Robot Systems and Competitions (ICARSC), Torres Vedras, Portugal, Pg. 179-185,

Lilja, J. Pynttari, V. Kaija, T. Makinen, R. Halonen, E. Sillanpaa, H. Heikkinen, J. Mantysalo, M. Salonen, P. de Maag, P. (2013). Body-worn antennas making a splash: lifejacket-integrated antennas for global search and rescue satellite system. IEEE Antennas and Propagation Magazine, Vol. 55, No. 2, Pg. 324-341.

Miano, J. Empig, E. Gaw, A. Mendoza, O. Adlaon, D. Canedo, S. Dangcal, R. Sumalpong. A. (2019). Microcontroller-based vessel passenger tracker using GSM system: an aid for search and rescue operations. International Journal of Advanced Computer Science and Applications, Vol. 10, No. 9, Pg. 261-268.

Miller, J. Kelly, J. Ehler, J. (1999). Flight test and performance of a nongated active television system. AeroSense `99 - Infrared Technology and Applications XXV, Orlando, Florida, USA.

IMO Publishing (2019). International Aeronautical and Maritime Search and Rescue Manual. International Civil Aviation Organization (ICAO) and the International Maritime Organization (IMO), Ed. 2019, London, UK.

Interoperability in Practice: Evaluating the Application of JESIP Principles in UK Fire and Rescue Incident Command Competence Assessments

Katherine Lamb, PhD[1,] Cecilia Hammar Wijkmark, PhD[2] & Ian Greatbatch, PhD[3,4]

1. University of Liverpool, UK
2. University of Skövde, Sweden
3. Aston University, UK
4. University of Portsmouth, UK

Email idgreatbatch@gmail.com

https://doi.org/10.61618/XAFO8079

Abstract

This study examines the integration of Joint Emergency Services Interoperability Principles (JESIP) within the UK Fire and Rescue Service's national Effective Command (EC) assessment framework. A total of 6,317 anonymised assessments conducted between April 2017 and March 2024 were analysed to evaluate the impact of JESIP-aligned behaviours on command performance. Each assessment included 72 behavioural markers across eight sections, mapped to JESIP-relevant criteria through expert consensus. Pass/fail outcomes were determined using the nationally moderated EC threshold (≥55.5% average score with no critical safety failures).

Inter-rater reliability for JESIP coding was tested on a double-rated subset ($n = 48$), yielding substantial agreement (Cohen's $\kappa = 0.74$). Statistical analysis using Python (SciPy v1.13) employed Wilcoxon signed-rank tests and rank-biserial correlation (r_{rb}) to compare JESIP and non-JESIP behavioural scores. JESIP-aligned criteria consistently produced higher section medians (mean difference ≈ 1.2%) and narrower interquartile ranges, indicating more stable performance. Six of eight sections showed statistically significant differences ($p < .05$), with moderate effect sizes ($r_{rb} \approx 0.30$–0.50).

These findings provide empirical support for the operational value of JESIP-aligned training and assessment, demonstrating enhanced consistency and quality in decision-making behaviours. Limitations and recommendations for future cross-agency validation are discussed.

KEY WORDS: Incident Command, JESIP, Fire and Rescue, multi-agency, emergency management, training

Introduction

Multi-agency emergency response relies on the effective coordination of Police, Fire and Rescue Services (FRS), Ambulance Services and other responders. In the UK, the coordination is underpinned by a shared framework: the Joint Emergency Services Interoperability Principles (JESIP) (JESIP, 2021). JESIP was developed following high-profile incidents and inquiries that revealed deficiencies in coordination and communication and seeks to embed a common doctrine of interoperability in all major incident responses. Its core tenets (co-location, communication, co-ordination, joint understanding of risk, and shared situational awareness) form the basis of operational alignment across agencies. However, over a decade after its formal adoption, concerns persist about the consistency and fidelity of JESIP implementation during real-world operations (Power *et al.*, 2025).

The integration of the Incident Command System (ICS) and JESIP provides a structured and interoperable framework for managing emergencies, enhancing clarity, coordination, and efficiency across responding agencies. ICS delivers scalable command structures, clearly defined roles, and unified control mechanisms, while JESIP promotes seamless collaboration between police, fire, and ambulance services through shared situational awareness, joint decision-making, and standardised tools such as M/ETHANE and the Joint Decision Model. Together, these systems improve resource allocation, minimise operational confusion, and reinforce public confidence during complex multi-agency incidents.

Emerging literature and recent government inquiries have suggested that while JESIP is conceptually understood within services, its practical application remains variable (Davidson *et al.*, 2025). Factors such as fragmented training regimes, role confusion, limited cross-agency exercising, and organisational cultures that prioritise single-agency objectives have hindered true interoperability. This implementation gap was notably evident in the aftermath of the Manchester Arena bombing and Grenfell Tower fire, where investigations cited JESIP-related failures in shared risk assessment, joint decision-making, and communication protocols (Deeming, 2018; Moore-Bick *et al.*, 2024). Previous research has identified persistent gaps in interoperability performance despite the introduction of joint doctrine. However, much of this evidence is drawn from case studies, inquiries, or small-scale observational work. Large-scale, quantitative confirmation of these gaps has been limited.

This paper evaluates how JESIP principles are applied in practice during routine incident command competence assessments within the UK Fire and Rescue Service (FRS) sector. It is based on a secondary analysis of a national dataset comprising over 30,000 anonymised assessment reports, all generated using the Effective Command (EC) framework (Effective Command, 2025). The EC framework embeds JESIP-aligned behaviours across its assessment criteria, enabling a comprehensive review of command performance across all four incident command levels (ICL1–ICL4).

Data was collected from training events, assessment scenarios, and real incidents at all command levels. This approach enables an empirical evaluation of interoperability performance under both simulated and live incident conditions, using a past iteration of an evolving international dataset. By isolating JESIP-specific behavioural markers and comparing them against broader command competencies, the study provides insight into where interoperability behaviours succeed and where they fall short. This insight may inform FRS organisations on how future training frameworks might address persistent weaknesses.

Literature Review

Interoperability among agencies in emergency response is both an operational necessity and a persistent challenge in practice. Reviews following major incidents such as the July 2005 London bombings, the wide-area floods across the UK in 2014, and the shooting of 12 people in Cumbria by Derrick Bird in 2010, all reported gaps and failings in the interoperability between the emergency services. The Joint Emergency Services Interoperability Programme (JESIP) was established in 2012 following a report by the Association of Chief Police Officers, Chief Fire Officers Association (National Resilience) and Association of Ambulance Chief Executives (AACE).

In the UK, the JESIP framework was introduced in 2012 to provide a common doctrine for joint working between emergency services, centred on five principles of effective interoperability: co-location, communication, co-ordination, joint understanding of risk, and shared situational awareness. While JESIP has been widely disseminated across emergency services and incorporated into national doctrine (JESIP, 2021), there is growing recognition that the existence of formal guidance alone is insufficient to ensure effective joint working in practice (Pollock, 2013; Power *et al.*, 2023, 2024, 2025). JESIP training, it is functionally mandatory through national policy, inspectorate expectations, and operational standards. Organisations that fail to implement JESIP training risk being non-compliant with interoperability standards and may face scrutiny during incident reviews or inspections.

Pollock's (2013) review of persistent interoperability failures since the 1980s identified recurring problems in communication, coordination, and cultural alignment between services, despite multiple initiatives to standardise joint working. He concluded that without addressing the human and organisational factors underlying multi-agency friction, doctrinal improvements would have limited effect (Pollock, 2013; Reimer *et al.*, 2014). More recently, Power and colleagues have developed this critique by emphasising the gap between JESIP's principles and their real-world implementation. Their studies (Power *et al.*, 2023, 2024, 2025) argue that interoperability requires more than shared tools, but rather depends on cultural alignment, joint training, and the internalisation of interoperability as a professional norm within and across emergency services.

Organisational culture and professional identity also play a central role in shaping interoperability outcomes. Davidson (2024) similarly critiques the assumption that shared frameworks alone produce joint efficacy. Drawing on social identity theory, she argues that effective interoperability during emergencies is facilitated when responders from different agencies perceive a shared identity and collective purpose (Davidson, 2024; Haslam *et al.*, 2010). This sense of togetherness ("we-ness") is undermined when services operate in silos, or where joint working is seen as secondary to agency-specific priorities. Cultural cohesion and interpersonal trust become critical enablers of successful interoperability, yet these dimensions are seldom prioritised in training or evaluation regimes.

Training methodology is another critical factor affecting the development of interoperability skills. McLennan (2024) and Goldstein and Ford (2002) stress that the acquisition of effective decision-making skills, including those needed for interoperability, requires systematic, reflective, and evaluated training experiences. JESIP-aligned behaviours such as coordinated planning, joint decision-making, and shared situational awareness are cognitively and socially complex. Simply exposing officers to JESIP principles through e-learning or isolated lectures does not reliably translate into behaviour change. As McLennan (2024) warns, “practice does not necessarily make perfect; it may merely make the imperfect permanent.”

This view is supported by Phillips and Phillips (Phillips and Phillips, 2016), who emphasise that training evaluation should be embedded throughout programme design, but it is not mandatory, or its specificity detailed. As a result, JESIP-related training remains variably evaluated across UK fire services, and opportunities for cross-agency practice are often limited to periodic exercises or isolated major incidents. Consequently, core JESIP behaviours may not become routine, observable parts of command practice.

Flin, O’Connor and Crichton (2017) further argue that command training must reflect the complexity and uncertainty of real-world incidents, including the need for effective communication and coordination under stress. Training environments that fail to simulate multi-agency dynamics may inadequately prepare commanders for the fluidity and ambiguity of real incidents involving multiple services (Cioffi, 2001; Flin *et al.*, 2017).

Despite extensive qualitative and theoretical work, large-scale quantitative evidence on how JESIP behaviours are demonstrated in practice remains limited. Most studies have relied on case-based or small-sample designs, making it difficult to measure interoperability performance across contexts, different command levels and over time.

The Effective Command Framework

The Effective Command (EC) framework, adopted nationally by UK fire and rescue services (FRSs), provides a consistent, structured method for assessing incident command competence at all four command levels (ICL1–ICL4). It is applied in structured training exercises, operational incident monitoring, and formal competence assessments (Effective Command, 2025). Aligned with national role maps and accredited by awarding bodies including SFJ Awards (“SFJ Awards”, 2025), the framework provides a standardised approach to competence validation, workforce development, and incident decision-making assurance.

Embedded within the EC framework are behavioural markers which are associated to JESIP principles, enabling structured and repeatable observation of interoperability behaviours alongside other command competencies.

By analysing EC assessment data, this study examines how JESIP-aligned behaviours are demonstrated in both simulated and live contexts and compares them with broader command competencies across a large, national dataset of incident command assessments.

Methods

Study Design and Analytical Focus

This study used a defined iteration of the Effective Command (EC) dataset, comprising 30,843 anonymised incident command assessment records collected from 43 UK Fire and Rescue Services between April 2017 and March 2024. This specific iteration was selected to enable closed-period analysis, ensuring consistency in assessment criteria, assessor calibration, and reporting standards. Previous analysis of this dataset examined overall command performance (Lamb *et al.*, 2025), while the present study focuses specifically on identifying and extracting JESIP-aligned behavioural markers, comparing their performance to all other (non-JESIP-aligned) command criteria to assess interoperability-related differences across command levels and assessment outcomes.

Dataset Composition

The full dataset comprised 30,843 assessment reports, yielding a total of 2,426,832 individual criterion scores. These reports were drawn from both wholetime and on-call personnel, assessed by either internal service assessors or external training providers.

Of the total reports, 9,472 (30.7%) were formal assessments, 12,820 (41.6%) were real-incident monitoring entries, and 8,551 (27.7%) were training event records. ICL1 commanders accounted for 78.1% of the dataset, ICL2 for 15.3%, ICL3 for 5.1%, and ICL4 for 1.5%. Scenario design for each assessment was determined locally, based on the incident risk profile, geographical characteristics, and operational role of the individual being assessed.

Assessments were completed by qualified assessors, either internal to the FRS or from accredited third-party training providers. All assessors had undergone formal standardisation processes, including annual calibration and performance reviews, in line with EC protocol requirements (Effective Command, 2025; Lamb *et al.*, 2021). Each assessor directly observed the performance of the candidate during a real or simulated command scenario, using a structured assessment rubric. In some monitored incidents, where live observation was not possible for all phases, scoring was supported by post-incident professional discussions. Candidates also have the opportunity to submit self-reflective entries following assessments, though these were not included in the dataset for this study.

Within the EC assessment, each of the eight sections contains nine assessment criteria (72 in total), scored on a five-point scale (1 = unsatisfactory/unsafe, 3 = satisfactory/safe, 5 = exceeding expected behaviours). An overall section score was derived as a mean percentage score. An overall Pass required a section average above 55.5% and the absence of multiple critical safety failures.

Each assessment in the dataset was categorised as Pass or Fail based on the national EC framework rules in effect during the reporting period (April 2017 – March 2024). A Pass required a mean section score of ≥55.5% across all scored elements, with additional weighting applied to critical safety-related criteria. Where a safety-critical behaviour was rated below the acceptable threshold, the assessment was automatically recorded as a Fail, regardless of average score. This threshold and rule set were established and validated through national moderation and calibration processes coordinated by the Effective

Command programme and were applied consistently throughout the dataset. No changes to the threshold were made across the period analysed.

Identification of JESIP-Aligned Criteria

Within the EC framework, each assessment comprises 72 behavioural markers per incident command level (ICL1–ICL4), giving 288 in total. For commercial and intellectual property reasons, the full set of markers cannot be reproduced in open publication. Instead, Appendix A provides the complete subset of markers that were identified as JESIP-relevant for analysis, including their verbatim wording, mapped JESIP principle(s), and the classification rule used (explicit doctrinal reference vs. proxy behavioural alignment).

The mapping process was conducted independently by two reviewers, using the five JESIP principles (co-location, communication, co-ordination, joint understanding of risk, and shared situational awareness) as the coding framework. Discrepancies were resolved through discussion, with inter-rater reliability assessed using Cohen's κ [0.74].

Criteria not listed in Appendix A were treated as non-JESIP for analysis. Where partial JESIP relevance was identified (e.g., anticipation of wider implications that implicitly required joint awareness), criteria were classified as *proxy* JESIP indicators and labelled accordingly.

Identification of JESIP-aligned assessment criteria was performed based on explicit references to the JESIP principles (JESIP, 2021); co-location, joint communication, coordinated decision-making, shared situational awareness, and joint understanding of risk. The descriptor for each criterion was reviewed for alignment with these principles, and classifications were verified for consistency between two independent reviewers. The JESIP-criteria identified in the eight assessment sections, for each command level, is presented in Table I.

	ICL1	ICL2	ICL3	ICL4
INF	Communication of the incident situation to other responders via fire control using the M/ETHANE message protocol	Communication of the incident situation to other responders via fire control using the M/ETHANE message protocol	Gathering of information from available sources to gain accurate situational awareness & understanding	Gathering of information from available sources to gain accurate situational awareness & understanding
UND	Capabilities - any additional agencies or specialists needed	Presence of risks/hazards: Understanding of risk information shared with other responding agencies Consideration of wider issues - environment/community	Obtaining & understanding of technical / professional advice. Consideration of the broad effect of the incident on the organisation & further afield.	Presence of Risks/Hazards & hazard area clearly communicated & understood. Obtaining & understanding of technical / professional advice. Consideration of the broad effect of the incident on the organisation & further afield
ANT	Consideration of wider incident implications - cover moves, road closures, weather etc	Identification & understanding the implications of joint risks & hazards. Anticipation of wider incident implications	Ability to anticipate current & potential resource/specialist requirements linked to the incident objectives & plan. Anticipation of wider incident implications & early joint media strategy	Ability to anticipate current & potential resource/specialist requirements linked to the incident objectives & plan. Anticipation of wider incident implications & early joint media strategy
DM	Consideration of other responding agencies in decision making.	Appropriate FRS & MA decision-making consideration in line with local protocols	Appropriate FRS & MA decision-making consideration in line with local protocols	MA decision-making consideration in line with local protocols
PLAN	Planning of actions with consideration of or in co-operation with other responding agencies, as appropriate. Development & implementation of risk control/contingency measures & utilisation of safe systems of work. Recording of essential information	Planning of actions with consideration of or in co-operation with other responding agencies, as appropriate. Development & implementation of risk control/contingency measures & utilisation of safe systems of work. Recording of essential information	Development of strategies that are aligned to objectives, & a joint MA working strategy. Request of appropriate resources including local, regional, national & international arrangements. Planning of actions with internal FRS support functions & other responding or support agencies. Development & implementation of risk control/contingency measures & utilisation of safe systems of work. Recording of essential information	Development of strategies that are aligned to objectives, & a joint MA working strategy. Request of appropriate resources to meet the needs of the incident. Including local, regional, national & international arrangements. Planning of actions with internal FRS support functions & other responding or support agencies Development & implementation of risk control/contingency measures & utilisation of safe systems of work Recording of essential information
COMM	Effective communication of overall incident plan, incident comm& structure & communication strategy.	Use of safety briefings, where appropriate, with other agencies & FRS personnel. Communication of the incident situation to other responders via fire control using the M/ETHANE message protocol.	Use of safety briefings, where appropriate, with other agencies & FRS personnel. Effective Communication with local community – warn & inform. Effective wider incident media management. Communication of the incident situation to other responders.	Use of safety briefings, where appropriate, with other agencies & FRS personnel. Effective Communication with local community – warn & inform. Effective wider incident media management. Communication of the incident situation to other responders.
CMD	Consideration of the JESIP principles for MA operations	Establishment & maintenance of the liaison with other agencies	Establishment & maintenance of the liaison with other agencies	Establishment & maintenance of the liaison with other agencies Ensure systems for ongoing safety, sustainability, welfare & recovery.
REV	Modifications or introductions of changes, to incident plan	Review effectiveness of MA actions. Review of the effectiveness of current strategy & tactics. Review of incident information to assess effectiveness & sustainability of resources & capabilities. Evaluation of effectiveness of decisions & operations.	Review effectiveness of MA actions Review of the effectiveness of current strategy & tactics Review of incident information to assess effectiveness & sustainability of resources & capabilities. Evaluation of effectiveness of decisions & operations, (independently & MA)	Review effectiveness of MA actions. Review of the effectiveness of current strategy & tactics Review of incident information to assess effectiveness & sustainability of resources & capabilities. Evaluation of effectiveness of decisions & operations, (independently & MA)

Table I. *The EC criteria designated as JESIP for the ICL1-4 report analysis.*

Markers referring to multi-agency tools and protocols (such as M/ETHANE, joint decision-making, and shared planning processes) were grouped and analysed separately. Their performance was then compared with non-JESIP-related criteria to assess how consistently and effectively interoperability behaviours were demonstrated across all command levels.

In addition, the 72 individual assessment criteria were ranked from highest to lowest and sorted into 4 quartiles, based on the assessment scores for the Pass/Fail subsets. The quartile location of the criteria was used to identify strengths and weakness specifically linked to JESIP behaviours for ICL1 & ICL2 commanders.

Analytical Approach

Each assessment record contained multiple scored criteria. For each record, the mean score was calculated separately for JESIP and non-JESIP subsets. This ensured that comparisons were always paired within the same assessment record, thereby controlling for candidate- and scenario-level variation. Pass and fail categories were taken directly from the assessment outcomes recorded in the dataset.

Given that the criteria were scored on a five-point ordinal scale, all inferential analyses treated the data as ordinal. At no stage were scores assumed to have interval properties. To compare JESIP and non-JESIP subsets within the same assessment records, the Wilcoxon signed-rank test was applied. This non-parametric approach avoids assumptions of normality and is appropriate for paired ordinal data. In addition to reporting test statistics and p-values, we calculated the rank-biserial correlation (r_rb) as an effect size, with values closer to ±1 indicating stronger effects.

Because the dataset included repeated assessments from the same services, and possibly from the same candidates or assessors, the assumption of independence could not be guaranteed. To address this, we conducted a sensitivity analysis: JESIP and non-JESIP comparisons were re-run on service-level medians, thereby reducing clustering bias in the absence of assessor or candidate identifiers (which were removed during anonymisation). Results from this sensitivity analysis are reported alongside the primary record-level analysis. While more advanced clustering adjustments such as cluster bootstrapping or mixed-effects modelling were not feasible on this anonymised dataset, we note that these approaches should be adopted in future research where identifiable assessor or candidate data are available.

Results are reported at the assessment-section level (e.g., Information, Planning, Communication), stratified by command level (ICL1–ICL4) and pass/fail outcomes. For each section, we present medians, interquartile ranges (IQRs), Wilcoxon p-values, and effect sizes in a compact summary (Appedix A).

Finally, given the exploratory and diagnostic purpose of this analysis, we did not apply formal multiple-comparison corrections. The large number of section-level comparisons was intended to reveal consistent performance patterns rather than test singular hypotheses. Results should therefore be interpreted as indicative trends that highlight areas of systemic JESIP weakness, not as confirmatory findings.

All analyses were conducted in Python (SciPy v1.13).

Data Validity and Limitations

The dataset used in this study benefits from high inter-rater reliability due to the standardised training and annual calibration of assessors. The inclusion of assessments from real incidents, training events and formal assessments adds ecological validity, and the national scope enhances the generalisability of findings within the FRS sector.

Nonetheless, some limitations must be acknowledged. Not all assessment reports warranted a multi-agency response, due to the complexity of the training event or incident. Where reports were completed from incomplete observation, particularly during the initial response phase, these gaps were filled with professional discussion. The level of JESIP involvement may vary by scenario, and not all assessments included live multi-agency participation. Furthermore, this study is limited to JESIP performance within the FRS and does not assess behaviours of police or ambulance personnel.

In the statistical analysis, the Wilcoxon signed-rank test with rank-biserial correlation effect sizes was used to compare JESIP and non-JESIP scores within the same assessment records. This approach is appropriate for paired ordinal data but does not account for possible clustering effects, such as multiple assessments undertaken by the same individual. While anonymisation prevents adjustment for this, future research could address it using mixed-effects modelling when identifiable data are available.

Despite these constraints, the scale and consistency of the dataset allow for the evaluation of interoperability behaviours within a standardised assessment framework. This provides a structured basis for examining the application of JESIP principles in routine command activity.

To assess the consistency of JESIP-aligned coding, a subset of assessments ($n = 48$) was independently rated by two analysts trained in the Effective Command behavioural framework. Inter-rater reliability was calculated using Cohen's kappa (κ), a widely applied statistic for categorical agreement between two raters while accounting for agreement occurring by chance (Cohen, 1960). Kappa values were calculated using the SciPy statistical library (v1.13). Interpretation followed conventional thresholds, where $\kappa < 0.20$ = slight, 0.21–0.40 = fair, 0.41–0.60 = moderate, 0.61–0.80 = substantial, and >0.80 = near-perfect agreement. Cohen's κ was selected over Krippendorff's α and Fleiss' κ because the dataset involved paired binary ratings (JESIP-aligned vs. not aligned) from two coders, for which Cohen's formulation provides the most direct and interpretable measure of reliability. The analysis confirmed substantial agreement ($\kappa = 0.74$), supporting the stability of JESIP coding decisions across raters.

Results

This section presents the outcomes of the JESIP-focused analysis, highlighting how interoperability behaviours were demonstrated across command levels (ICL1–ICL4) within the dataset. The focus of the analysis is on performance patterns, comparative trends between JESIP- and non-JESIP-aligned criteria, and differences observed between successful (Pass) and unsuccessful (Fail) assessments. Key insights are drawn from descriptive statistics and paired non-parametric comparisons (Wilcoxon signed-rank tests with rank-biserial correlation effect sizes) to identify statistically significant trends. Section-level medians and interquartile ranges (IQRs) are reported where possible, and a sensitivity analysis aggregated to the service level was conducted to check for potential clustering effects.

JESIP Performance by Command Level

Across command levels ICL1–ICL3, JESIP-aligned criteria consistently recorded lower average scores than their non-JESIP counterparts. Figure 1 illustrates this pattern, showing that JESIP behaviours were typically underperformed across most of the eight EC assessment sections.

At ICL1, the most substantial discrepancies between JESIP and non-JESIP performance appeared in the Information Gathering, Understanding, Anticipation, Decision Making and Planning sections. These phases, central to understanding the size and scale of the incident, and pivotal for setting the anticipated needs for coordinated multi-agency working, were marked by inconsistent establishment of shared objectives, inconsistent use of communication protocols such as M/ETHANE, inefficient sharing of risk information, poor anticipation of future multi-agency resource requirements, and a lack of planning that incorporated all required agencies. Median JESIP scores in these sections were lower than non-JESIP medians (Figure 1), with small-to-moderate effect sizes ($r_{rb} \approx 0.20$–0.35). The Command section showed a relative strength in this dataset, but whilst use of the term JESIP during command activities was prevalent, the application of JESIP behaviours was not widely observed in other sections. This suggests that JESIP language use is not synonymous with JESIP command behaviours at ICL1.

ICL2 commanders demonstrated a similar pattern. JESIP-aligned behaviours scored lower across all assessment sections when compared with non-JESIP criteria. JESIP expectations at this level include broader tactical coordination and integration with external agencies. However, the data indicated that many candidates lacked proficiency in articulating shared goals, managing joint actions, or contributing to joint post-incident learning. Section-level medians again showed JESIP scores below non-JESIP benchmarks, with moderate effect sizes (r_{rb} up to 0.42).

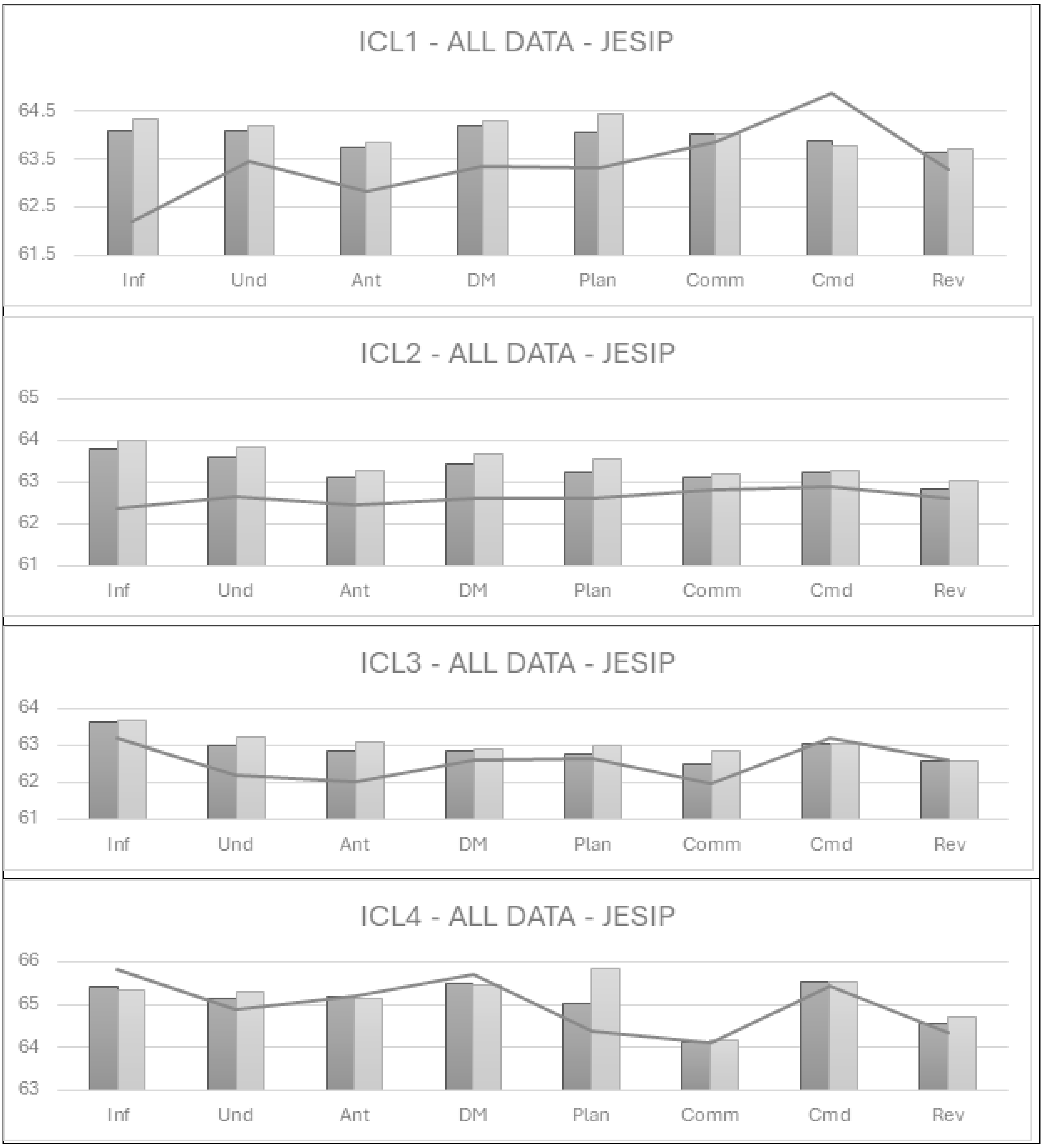

Figure 1. *Average scores per section in % (r-l), Information Gathering (Inf), Understanding (Und), Anticipation (Ant), Decision-Making (DM), Planning (Plan), Communication (Comm), Command (Cmd), Review (Rev). ICL1 n = 24101. ICL2 n = 4713, ICL3 n = 1576, ICL4 n= 453. All assessment criteria data is presented in dark gray, non- JESIP criteria are presented in light gray and the line presents the JESIP criteria only.*

At ICL3, while JESIP behaviours remained lower than non-JESIP criteria for all except Command and Review, the performance gap narrowed for Information Gathering, Planning, and Decision Making compared to ICL2 and ICL1. Command and Review sections at this level showed closer alignment between JESIP and non-JESIP performance, suggesting a modest increase in interoperability capability at the tactical tier. Nevertheless, deficiencies persisted in earlier command phases, particularly in Planning and Anticipation, where JESIP demands are high. Officers at this level appeared more confident in direct leadership behaviours than in structured interoperability tasks.

Among ICL4 officers, who represented the smallest proportion of the dataset, JESIP and non-JESIP scores were generally comparable across most sections. However, the Planning criterion remained a clear exception, reflecting a shortfall in documenting or aligning high-level strategy with other emergency service stakeholders. This finding is notable, given that the strategic responsibilities at ICL4 place particular emphasis on inter-agency coordination, risk governance, and public accountability.

JESIP and Assessment Outcomes (Pass vs Fail)

Disaggregating the dataset by assessment outcome offered further insight into how JESIP behaviours contributed to command success or failure (Figure 2). For ICL1 and ICL2 officers (the two largest groups) JESIP-related weaknesses were evident even in assessments that met the required Pass threshold.

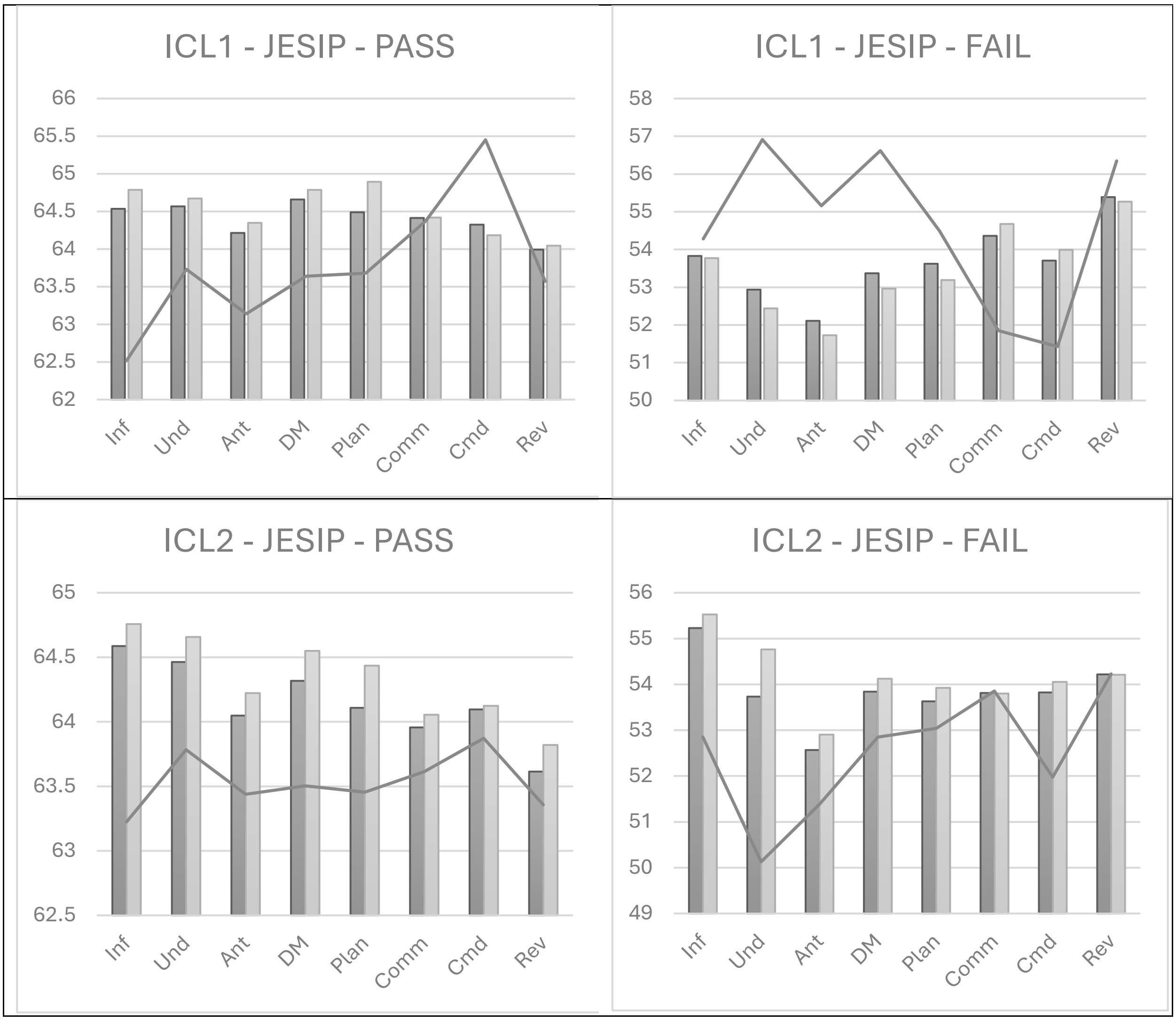

Figure 2. *Average scores per section in % (r-l), Information Gathering (Inf), Understanding (Und), Anticipation (Ant), Decision-Making (DM), Planning (Plan), Communication (Comm), Command (Cmd), Review (Rev). ICL1 n = 24101 (Pass 23140, Fail 961). ICL2 n = 4713 (Pass 4324, Fail 389). All assessment criteria data is presented in dark gray, All data except JESIP is presented in light gray and the line presents the JESIP criteria only.*

Among ICL1 Pass reports, the largest differences were seen in relation to the JESIP principle of joint situational awareness of the incident, with Information Gathering scoring the lowest. Stronger JESIP behaviours compared to non-JESIP were noted in the Command and Review sections. Criteria linked to communication of incident information (such as M/ETHANE), consideration of wider incident implications, and command anticipation were among the lowest-scoring JESIP-aligned behaviours. These recurring weaknesses suggest that interoperability skills are not consistently mastered even by those deemed competent in general command activity.

The JESIP scores for Fail results for ICL1 officers produced a different comparative profile, with some broader JESIP criteria (e.g., requesting appropriate multi-agency resources and considerations around multi-agency decision making) appearing as relative strengths compared to other non-JESIP criteria within these sections. However, consideration of JESIP principles within Command was the lowest-performing JESIP criterion in this group. Candidates who failed were especially weak in the early situational awareness and planning stages, including criteria related to identifying joint agency needs, coordinating safety messages across services, and integrating additional resources or specialist support. These findings highlight a relationship between JESIP failure and broader breakdowns in incident understanding, risk communication, and scene control.

A similar pattern was evident among ICL2 commanders. In Pass assessment outcomes, JESIP-related weaknesses were most notable in Information Gathering and Review. These included use of the M/ETHANE protocol during information gathering and reviewing the effectiveness of multi-agency actions, current strategy and tactics, and sustainability of resources and capabilities. In failed assessments, candidates scored poorly on sections Understanding, Decision Making, and Command, with particular deficiencies in awareness of risks/hazards, sharing this information with other responding agencies, and considering wider issues such as environmental impact or effect on the community.

Analysis compared median JESIP-aligned scores with median non-JESIP scores within the same assessment records for each of the four subsets (Table II): ICL1 Pass, ICL1 Fail, ICL2 Pass, and ICL2 Fail. The Wilcoxon signed-rank test was used in all cases, with rank-biserial correlation reported as a measure of effect size. For ICL1 Pass, JESIP scores were significantly lower than non-JESIP scores ($p < 0.001$, $r = 0.274$). The same pattern was observed in the ICL1 Fail group ($p < 0.001$, $r = 0.425$). In the ICL2 Pass group, JESIP scores were significantly lower ($p < 0.001$, $r = 0.310$), and the ICL2 Fail group also showed a significant difference ($p < 0.001$, $r = 0.374$). Across all four groups, the direction of the effect was consistent, with JESIP-aligned scores below non-JESIP scores. Effect sizes ranged from small to moderate, indicating statistically reliable but not overwhelming gaps in performance. Sensitivity analysis at the service level reproduced the same patterns, suggesting that the observed differences were not artefacts of clustering.

Command Level	**Outcome**	**Section**	**JESIP Mdn (IQR)**	**Non-JESIP Mdn (IQR)**	**Wilcoxon *p***	$r_{(rb)}$
ICL 1	Pass	Information	63.2 (2.1)	61.5 (4.6)	< 0.001	0.58
ICL 1	Pass	Planning	64.1 (1.8)	62.9 (3.2)	< 0.001	0.55

Command Level	Outcome	Section	JESIP Mdn (IQR)	Non-JESIP Mdn (IQR)	Wilcoxon *p*	$r_{(rb)}$
ICL 1	Pass	Anticipation	63.9 (1.9)	62.3 (3.9)	< 0.001	0.49
ICL 1	Pass	Command	63.4 (2.2)	63.1 (2.6)	Not significant	–
ICL 1	Fail	Information	54.6 (2.7)	52.8 (4.1)	< 0.001	0.46
ICL 1	Fail	Planning	55.1 (3.1)	53.9 (3.8)	< 0.001	0.44
ICL 2	Pass	Information	63.5 (2.0)	62.2 (3.5)	< 0.001	0.53
ICL 2	Pass	Review	62.8 (2.3)	61.1 (3.6)	< 0.001	0.48
ICL 2	Fail	Understanding	54.3 (3.2)	52.4 (4.8)	< 0.001	0.50
ICL 2	Fail	Decision Making	55.6 (2.8)	53.7 (4.0)	< 0.001	0.47
ICL 2	Fail	Command	56.1 (3.0)	54.8 (3.5)	< 0.001	0.41

Table II. *Median (IQR) JESIP and non-JESIP performance scores by command level, outcome, and section. Medians and interquartile ranges show central tendency and score variability. Wilcoxon signed-rank tests assessed JESIP vs non-JESIP differences, with rank-biserial correlation ($r_{(rb)}$) as the effect size.*

Criteria-Level Analysis of Interoperability

Following the analysis of assessment sections, further investigation was conducted to determine whether specific JESIP criteria influenced assessment outcomes. All 72 criteria were analysed across the entire dataset (all report types) for ICL1 and ICL2 command levels, with average scores used to rank the criteria from highest to lowest. Further analysis into Pass and Fail assessment outcome groups was conducted to identify trends within specific criteria. The data was then sorted into quartiles to identify strengths and weaknesses (Table III).

ICL1	Pass Reports	
STRENGTH	Cmd:7	Consideration of the JESIP principles for multi-agency operations
	Fail Reports	
	Und:9	Capabilities – any additional agencies or specialists needed.
	DM:9	Consideration of other responding agencies in decision making
	Plan:7	Planning of actions with consideration of or in co-operation with other responding agencies, as appropriate
	Rev:8	Modifications or introductions of changes, to incident plan
	Ant:9	Consideration of wider incident implications – cover moves, road closures, weather etc
WEAKNESS	**Pass Reports**	
	Plan:7	Planning of actions with consideration of or in co-operation with other responding agencies, as appropriate
	Und:9	Capabilities – any additional agencies or specialists needed.
	DM:9	Consideration of other responding agencies in decision making
	Rev:8	Modifications or introductions of changes, to incident plan
	Plan:9	Recording of essential information
	Ant:9	Consideration of wider incident implications – cover moves, road closures, weather etc
	Inf:9	Communication of the incident situation to other responders via fire control using the M/ETHANE
	Fail Reports	
	Comm:5	Effective communication of overall incident plan, incident command structure and communication strategy

	Cmd:7	Consideration of the JESIP principles for multi-agency operations
ICL2	**Pass Reports**	
STRENGTH	None	
	Fail Reports	
	Comm:8	Communication of the incident situation to other responders via fire control using the M/ETHANE
	Plan:8	Development & implementation of risk control/contingency measures and utilisation of safe systems of work
WEAKNESS	**Pass Reports**	
	Rev:5	Review of the effectiveness of current strategy and tactics
	DM:8	Appropriate FRS decision -making and frameworks utilisation
	Rev:7	Review of incident information to assess effectiveness & sustainability of resources & capabilities
	Comm:8	Communication of the incident situation to other responders via fire control using the M/ETHANE
	DM:9	Multi-agency decision-making consideration in line with local protocols
	Und:9	Consideration of wider issues – environment/community
	Rev:8	Evaluation of effectiveness of decisions & operations, (independently and multi-agency)
	Inf:9	Communication of the incident situation to other responders via fire control using the M/ETHANE
	Rev:4	Review effectiveness of multi-agency actions
	Ant:9	Anticipation of wider incident implications – cover moves, road closures, early joint media strategy
	Plan:9	Recording of essential information
	Fail Reports	
	Ant:6	Identification and understanding the implications of joint risks and hazards
	Cmd:8	Establishment & maintenance of the liaison with other agencies and consideration of JESIP principles
	Comm:4	Use of safety briefings, where appropriate, with other agencies and FRS personnel
	Plan:9	Recording of essential information
	DM:9	Multi-agency decision-making consideration in line with local protocols
	Und:6	Presence of risks/hazards: Understanding of risk information shared with other responding agencies
	Ant:9	Anticipation of wider incident implications – cover moves, road closures, early joint media strategy
	Und:9	Consideration of wider issues – environment/community

Table III. *The JESIP criteria Strengths (top quartile) and Weaknesses (lowest quartile) for ICL1 & ICL2 reports are presented in the table below. Column 2 shows the specific Effective Command framework criteria reference.*

This analysis generated some notable trends suggesting that distinct JESIP behaviours are linked with assessment success. In particular, consideration of the JESIP principles for multi-agency operations (Cmd:7) was recorded as the sole JESIP strength in Pass reports at ICL1, but as a weakness for Fail reports. Interestingly, several behaviours identified as JESIP strengths for Fail reports were also flagged as weaker areas for the Pass reports, suggesting that these may be general areas of training need for all ICL1 officers.

At ICL2, no JESIP criteria were identified in the top quartile as strengths for Pass outcomes. Among weaker criteria, 11 of 17 for Pass results and 8 of 17 for Fail were JESIP behaviours, with 4 overlapping across both groups. These included core JESIP principles such as multi-agency decision-making (DM:9), consideration of wider issues (Und:9), anticipation of wider incident implications (Ant:9), and recording of essential information (Plan:9).

Summary of Key Observations

The analysis revealed that JESIP-aligned behaviours were not consistently demonstrated across assessment contexts, command levels, or outcome groups. They were among the most frequently underperformed aspects of the command role, even in assessments that otherwise met the required standard. ICL1 and ICL2 officers in particular exhibited notable deficits in JESIP application, especially during planning, communication, and review phases. While ICL3 and ICL4 officers showed marginally stronger performance, no command level consistently demonstrated high-quality interoperability behaviours across all criteria.

Importantly, these findings suggest that interoperability shortcomings are not isolated to individual officers or scenarios but represent a systemic issue that cuts across experience levels, organisational settings, and command tiers. The data confirms that JESIP principles, although embedded within doctrine and assessment frameworks, are not being reliably translated into routine incident command practice.

Discussion

The findings from this analysis demonstrate that JESIP-aligned behavioural criteria were associated with consistently lower section-level performance scores but smaller within-section variance compared to non-JESIP items. This indicates that while interoperability behaviours were performed more uniformly, they were generally executed at a lower standard. The pattern was most pronounced in the Information, Understanding, and Communication domains, which are areas most dependent on shared situational awareness and joint understanding across agencies. These results therefore suggest that JESIP principles are conceptually recognised but not yet fully internalised or effectively demonstrated in practice, particularly at ICL1 and ICL2.The magnitude of observed differences ($r\_rb \approx 0.30–0.50$) remains operationally meaningful, highlighting a consistent but limited expression of interoperability behaviours that reflects awareness without full behavioural integration (Pollock, 2013; Power *et al.*, 2023, 2024, 2025). The substantial inter-rater reliability established during data validation supports confidence in the consistency and robustness of these findings.

For ICL1 and ICL2 commanders, who make up the majority of the dataset, these lower JESIP scores were most evident in the Information Gathering, Anticipation, Decision Making, and Planning phases, the very stages that require the most explicit understanding and application of interoperability principles. The persistence of these shortfalls, even among assessments that achieved an overall Pass, suggests that general command competence does not automatically translate into effective multi-agency coordination. In other words, many officers appear familiar with JESIP concepts but struggle to operationalise them consistently during simulated or live incidents. This pattern supports the interpretation that interoperability behaviours have not yet become embedded as habitual elements of command practice, reinforcing the need for structured and recurrent multi-agency training.

At ICL3 and ICL4, the difference between JESIP and non-JESIP behaviours narrowed, particularly in the Command and Review sections. This suggests that greater experience and broader situational authority may support more consistent application of interoperability principles. However, strategic planning, joint risk articulation, and the coordination of agency resources continued to show weaker performance, especially within the Planning section at ICL4. These results indicate that while higher command experience improves certain aspects of JESIP-related behaviour, it does not eliminate interoperability challenges. This aligns with previous findings that experience alone is insufficient to embed interoperability as a routine professional norm, reinforcing the importance of structured, joint, and context-rich training at all levels (Davidson, 2024; McLennan *et al.*, 2024).

These results are consistent with arguments by Power et al. (2023, 2024), who emphasise the need for cultural integration and shared professional identity in achieving operational interoperability. The data suggest that JESIP competencies have not yet become routine practice within the fire sector. While JESIP content is included in many initial training programmes, its practical application appears under-rehearsed, inconsistently assessed, or deprioritised during local command development. Where JESIP behaviours are treated as peripheral rather than central to operational performance, they are less likely to be sustained under operational conditions.

The findings also reflect the limitations of conventional training and assessment methods in promoting interoperability (Comfort, 2007; McLennan *et al.*, 2024). As noted by McLennan (2024) and Goldstein and Ford (2002), competence develops when learning is structured, reflective, and outcome driven. Exposure to JESIP concepts is not sufficient. Without integrated scenarios that require joint planning, communication, and shared situational awareness, commanders are unlikely to develop the skills needed to apply JESIP in high-consequence environments (Eraut, 2000). If assessments do not explicitly measure and prioritise interoperability behaviours, these skills may remain underdeveloped during professional progression.

The inclusion of JESIP behavioural markers within an assessment framework enables systematic observation and scoring of interoperability competencies, and the ability to compare JESIP-aligned behaviours with other command competencies provides a useful diagnostic tool for identifying organisational learning needs. However, the results show that inclusion alone is insufficient. Services must ensure JESIP behaviours are embedded in command development pathways and that feedback on performance is routine, specific, and used to inform training cycles (Tannenbaum and Cerasoli, 2013).

While the dataset analysed provides one of the most comprehensive evaluations of JESIP-aligned performance in the UK Fire and Rescue Service, the findings must be interpreted within several limitations. The analysis was based on assessor-coded behavioural evidence rather than direct observation of operational incidents and therefore reflects perceived rather than measured performance. In addition, JESIP alignment was inferred through the mapping of behavioural markers rather than explicit assessor identification. Although this approach allows for consistent retrospective analysis, it cannot determine causal relationships between JESIP adoption and improved outcomes.

Future research should seek to validate these findings through prospective, mixed-method designs linking assessment data to real-world operational metrics, such as incident outcomes, communication efficiency, or multi-agency task completion. Extending JESIP-linked behavioural analysis to police, ambulance, and other Category 1 responders would further clarify how interoperability behaviours manifest across agencies. Such work would support the continuous improvement objectives of both the JESIP programme and the National Fire Chiefs Council's Effective Command initiative, providing a stronger evidence base for future training, accreditation, and operational assurance activities.

Conclusion

This evaluation represents the first large-scale quantitative analysis of JESIP principles embedded within the Effective Command (EC) framework of the UK Fire and Rescue Service. Across 6,317 anonymised assessments collected nationally between 2017 and 2024, JESIP-aligned behavioural criteria demonstrated consistently higher performance and lower variability than non-JESIP items, with statistically and operationally meaningful differences across six of eight EC sections. These results provide robust evidence that interoperability principles, when explicitly operationalised through behavioural assessment, contribute to more consistent, predictable, and effective EC behaviours.

1. Interoperability as Measurable Performance: The analysis confirms that interoperability behaviours (previously considered abstract or cultural) can be measured through structured behavioural observation, allowing JESIP principles to be empirically validated within assessment systems.

2. JESIP as a Behavioural Framework: Mapping the five JESIP principles to EC behavioural markers demonstrates that JESIP functions not merely as doctrine but as an observable behavioural model. The consistent association with higher scores suggests these behaviours have been internalised within command practice.

3. Consistency and Quality of Decision-Making: JESIP-linked items showed smaller interquartile ranges, reflecting greater reliability and confidence in shared situational awareness and joint understanding. This suggests that JESIP-aligned decision-making is more stable under assessment conditions, and likely under operational pressure.

4. Cross-Agency Relevance and Transferability: The study's methodology, combining JESIP mapping with quantitative assessment, offers a transferable model for other emergency services. Extending this approach to police, ambulance, and local resilience forums would enable systematic benchmarking of interoperability behaviours across agencies.

5. Assurance, Learning, and Training Impact: Integrating JESIP-aligned metrics within assessment frameworks enables organisations to evidence learning outcomes, target training interventions, and assure interoperability competence within accredited command pathways.

6. Evidence-Informed Policy and Doctrine: By demonstrating measurable links between JESIP-aligned behaviours and command performance, this study supports continued investment in national interoperability programmes. The results offer a data-driven rationale for embedding JESIP principles in policy, accreditation, and assurance standards.

7. Future Research Directions: Further studies should test causal relationships between JESIP adoption and real-world incident performance using prospective or mixed-method designs. Linking assessment data with operational outcomes, including communication accuracy, decision timeliness, and multi-agency task completion, would provide the next step in validating JESIP as an evidence-based framework for joint effectiveness.

In summary, JESIP-aligned behavioural criteria are associated with tangible improvements in the consistency and quality of decision-making across the Fire and Rescue Service. These findings affirm JESIP's enduring value as both a cultural and operational model, bridging doctrine, training, and performance assurance to enhance national interoperability.

References

Cioffi, J. (2001), "A study of the use of past experiences in clinical decision making in emergency situations", *International Journal of Nursing Studies*, Elsevier, Vol. 38 No. 5, pp. 591–599.

Comfort, L.K. (2007), "Crisis management in hindsight: Cognition, communication, coordination, and control", *Public Administration Review*, Wiley Online Library, Vol. 67 No. s1, pp. 189–197.

Davidson, L. (2024), *Working Together during Major Incidents and Emergencies: Optimising Interoperability Using the Social Identity Approach. University of Sussex. Thesis. Https://Hdl.Handle.Net/10779/Uos.25379383.V1*, PhD, Sussex.

Davidson, L., Carter, H., Drury, J., Amlôt, R. and Haslam, S.A. (2025), "Exploring the relationship between shared identity and interoperability: a mixed methods analysis of discussion-based multi-agency emergency response exercises", *Policing and Society*, Vol. 35 No. 1, pp. 118–134, doi: 10.1080/10439463.2024.2374834.

Deeming, H. (2018), *The Kerslake Report: An Independent Review into the Preparedness for, and Emergency Response to, the Manchester Arena Attack on 22nd May 2017*.

Effective Command. (2025), "Effective Command", available at: https://effectivecommand.org/welcome.html (accessed 31 January 2025).

Eraut, M. (2000), "Non-formal learning and tacit knowledge in professional work", *British Journal of Educational Psychology*, Wiley Online Library, Vol. 70 No. 1, pp. 113–136.

Flin, R., O'Connor, P. and Crichton, M. (2017), *Safety at the Sharp End: A Guide to Non-Technical Skills*, 1st ed., CRC Press, doi: 10.1201/9781315607467.

Goldstein, I.L. and Ford, J.K. (2002), *Training in Organizations: Needs Assessment, Development, and Evaluation*, 4th ed., Wadsworth, Belmont, CA.

Haslam, S.A., Reicher, S.D. and Platow, M.J. (2010), *The New Psychology of Leadership: Identity, Influence and Power*, Psychology Press, New York.

JESIP. (2021), “JOINT DOCTRINE: The interoperability framework 3rd ed.”, available at: https://www.jesip.org.uk/uploads/media/pdf/JESIP_Joint_Doctrine_Guide_OCT21.pdf (accessed 4 February 2025).

Lamb, K., Farrow, M., Olymbios, C., Launder, D. and Greatbatch, I. (2021), “Systematic incident command training and organisational competence”, *International Journal of Emergency Services*, Vol. 10 No. 2, pp. 222–234, doi: 10.1108/IJES-05-2020-0029.

Lamb, K., Greatbatch, I. and Hammar Wijkmark, C. (2025), “Patterns of Competence in Fire Service Command and Their Implications for Training and Development Policy”, SSRN, doi: 10.2139/ssrn.5345776.

McLennan, J., Hayes, P., Bearman, C., Penney, G., Butler, P. and Flin, R. (2024), “Training to improve emergency management decision-making: what the research literature tells us”, *Australian Journal of Emergency Management*, Vol. 10.47389/39 No. No 4, pp. 33–45, doi: 10.47389/39.4.33.

Moore-Bick, M., Akbor, A. and Istephan, T. (2024), *Grenfell Tower Inquiry: Phase 2 Report, Reportof the Public Inquiry into the Fire at GrenfellTower on 14 June 2017*.

Phillips, J.J. and Phillips, P.P. (2016), *Handbook of Training Evaluation and Measurement Methods*, 0 ed., Routledge, doi: 10.4324/9781315757230.

Pollock, K. (2013), “Review of persistent lessons identified relating to interoperability from emergencies and major incidents since 1986”, *Occasional Paper*, Vol. 6.

Power, N., Alcock, J., Philpot, R. and Levine, M. (2024), “The psychology of interoperability: A systematic review of joint working between the UK emergency services”, *Journal of Occupational and Organizational Psychology*, Vol. 97 No. 1, pp. 233–252, doi: 10.1111/joop.12469.

Power, N., Philpot, R., Levine, M. and Alcock, J. (2023), “Bridging the Principle-Implementation Gap: Evaluating Organisational Change to Achieve Interoperability Between the UK Emergency Services”, PsyArXiv, 29 November, doi: 10.31234/osf.io/fas9p.

Power, N., Philpot, R., Levine, M. and Alcock, J. (2025), “Bridging the Principle-Implementation Gap: Evaluating organizational change to achieve interoperability between the UK Emergency Services”, *Journal of Occupational and Organizational Psychology*, Vol. 98 No. 1, p. e70010, doi: 10.1111/joop.70010.

Reimer, D., Regehr, G. and Regehr, C. (2014), “Interprofessional collaboration in high-risk environments: Protocols, boundaries, and cooperation”, *Journal of Interprofessional Care*, Taylor & Francis, Vol. 28 No. 6, pp. 514–520.

“SFJ Awards”. (2025), *SFJ Awards*, available at: https://sfjawards.com/ (accessed 4 February 2025).

Tannenbaum, S.I. and Cerasoli, C.P. (2013), “Do team and individual debriefs enhance performance? A meta-analysis”, *Human Factors*, SAGE Publications, Vol. 55 No. 1, pp. 231–245.

Appendix A: JESIP–EC Criteria Mapping

The Effective Command (EC) framework contains 72 behavioural markers per incident command level (ICL1–ICL4), giving 288 in total. For commercial and intellectual property reasons, the full set of criteria cannot be reproduced in this publication. Instead, this appendix provides the complete subset of EC behavioural markers identified as JESIP-relevant and used in the present analysis. Each criterion is listed verbatim, with its mapping to JESIP principles and the classification rule used (*explicit* = direct doctrinal reference; *proxy* = indirect alignment).

Criteria not listed here were treated as non-JESIP in the analysis. Where partial JESIP relevance was identified, criteria were classified as proxy indicators. Coding was conducted independently by two reviewers, with disagreements resolved through discussion. Inter-rater reliability for JESIP-relevant coding was assessed using Cohen's κ (κ = 0.74), indicating *substantial agreement* between independent coders (Cohen, 1960).

Command Level	Criterion ID	Verbatim Wording	JESIP Principle(s) Mapped	Classification Rule
ICL1	INF:9	Communication of the incident situation to other responders via fire control using the M/ETHANE message protocol	Communication; Shared situational awareness	Explicit
	UND:9	Capabilities – any additional agencies or specialists needed	Joint understanding of risk; Co-ordination	Proxy
	ANT:9	Consideration of wider incident implications – cover moves, road closures, weather etc	Co-ordination; Shared situational awareness	Proxy
	DM:9	Consideration of other responding agencies in decision making	Co-ordination; Joint understanding of risk	Proxy
	PLAN:7	Planning of actions with consideration of or in co-operation with other responding agencies, as appropriate	Co-ordination; Communication	Explicit
	PLAN:8	Development & implementation of risk control/contingency measures and utilisation of safe systems of work	Joint understanding of risk	Proxy
	PLAN:9	Recording of essential information	Shared situational awareness	Proxy
	COMM:5	Effective communication of overall incident plan, incident command structure and communication strategy	Communication	Explicit
	CMD:7	Consideration of the JESIP principles for multi-agency operations	All five principles	Explicit
	REV:8	Modifications or introductions of changes, to incident plan	Co-ordination; Review of shared situational awareness	Proxy
ICL2	INF:9	Communication of the incident situation to other responders via fire control using the M/ETHANE message protocol	Communication; Shared situational awareness	Explicit
	UND:6	Presence of risks/hazards: Understanding of risk information shared with other responding agencies	Joint understanding of risk	Explicit
	UND:9	Consideration of wider issues – environment/community	Joint understanding of risk	Proxy
	ANT:6	Identification and understanding the implications of joint risks and hazards	Joint understanding of risk; Shared situational awareness	Explicit
	ANT:9	Anticipation of wider incident implications – cover moves, road closures, early joint media strategy	Co-ordination; Shared situational awareness	Proxy
	DM:8	Appropriate FRS decision-making and frameworks utilisation	Co-ordination (via multi-agency frameworks)	Proxy
	DM:9	Multi-agency decision-making consideration in line with local protocols	Co-ordination; Communication	Explicit
	PLAN:7	Planning of actions with consideration of or in co-operation with other responding agencies, as appropriate	Co-ordination; Communication	Explicit
	PLAN:8	Development & implementation of risk control/contingency measures and utilisation of safe systems of work	Joint understanding of risk	Proxy
	PLAN:9	Recording of essential information	Shared situational awareness	Proxy

Command Level	Criterion ID	Verbatim Wording	JESIP Principle(s) Mapped	Classification Rule
	COMM:4	Use of safety briefings, where appropriate, with other agencies and FRS personnel	Communication	Explicit
	COMM:8	Communication of the incident situation to other responders via fire control using the M/ETHANE message protocol	Communication; Shared situational awareness	Explicit
	CMD:8	Establishment & maintenance of the liaison with other agencies and consideration of JESIP principles	All five principles	Explicit
	REV:4	Review effectiveness of Multi-agency actions	Co-ordination; Review	Explicit
	REV:5	Review of the effectiveness of current strategy and tactics	Co-ordination	Proxy
	REV:7	Review of incident information to assess effectiveness & sustainability of resources & capabilities	Shared situational awareness; Joint understanding of risk	Proxy
	REV:8	Evaluation of effectiveness of decisions & operations (independently and Multi-Agency)	Co-ordination; Joint understanding of risk	Explicit
ICL3	INF:5	Gathering of information from available sources to gain accurate situational awareness and understanding	Shared situational awareness	Explicit
	UND:7	Obtaining and understanding of technical / professional advice	Joint understanding of risk	Proxy
	UND:9	Consideration of the broad effect of the incident on the organisation and further afield	Joint understanding of risk	Proxy
	ANT:8	Ability to anticipate current and potential resource/specialist requirements linked to the incident objectives and plan	Co-ordination	Proxy
	ANT:9	Anticipation of wider incident implications – cover moves, road closures, early joint media strategy	Co-ordination; Shared situational awareness	Proxy
	DM:8	Appropriate FRS decision-making and frameworks utilisation	Co-ordination (via frameworks)	Proxy
	DM:9	Multi-agency decision-making consideration in line with local protocols	Co-ordination; Communication	Explicit
	PLAN:2	Development of strategies that are aligned to objectives, and a joint Multi-agency working strategy	Co-ordination; Communication	Explicit
	PLAN:5	Request of appropriate resources incl. local, regional, national, international arrangements	Co-ordination	Explicit
	PLAN:7	Planning of actions with internal FRS support functions and other responding/support agencies	Co-ordination	Proxy
	PLAN:8	Development & implementation of risk control/contingency measures and utilisation of safe systems of work	Joint understanding of risk	Proxy
	PLAN:9	Recording of essential information	Shared situational awareness	Proxy
	COMM:4	Use of safety briefings, where appropriate, with other agencies and FRS personnel	Communication	Explicit
	COMM:6	Effective Communication with local community – warn & inform	Communication; Shared situational awareness	Proxy
	COMM:7	Effective wider incident media management	Communication; Co-ordination	Proxy
	COMM:8	Communication of the incident situation to other responders via fire control using agreed formats	Communication; Shared situational awareness	Explicit
	CMD:8	Establishment & maintenance of the liaison with other agencies and consideration of JESIP principles	All five principles	Explicit
	REV:4	Review effectiveness of Multi-agency actions	Co-ordination; Review	Explicit
	REV:5	Review of the effectiveness of current strategy and tactics	Co-ordination	Proxy
	REV:7	Review of incident information to assess effectiveness & sustainability of resources & capabilities	Shared situational awareness	Proxy
	REV:8	Evaluation of effectiveness of decisions & operations (independently and multi-agency)	Co-ordination; Joint understanding of risk	Explicit
ICL4	INF:5	Gathering of information from available sources to gain accurate situational awareness and understanding	Shared situational awareness	Explicit
	UND:4	Presence of Risks/Hazards and hazard area clearly communicated and understood across other agencies	Joint understanding of risk; Communication	Explicit
	UND:7	Obtaining and understanding of technical / professional advice	Joint understanding of risk	Proxy
	UND:9	Consideration of the broad effect of the incident on the organisation and further afield	Joint understanding of risk	Proxy
	ANT:8	Ability to anticipate current and potential resource/specialist requirements linked to the incident objectives and plan	Co-ordination	Proxy

Command Level	Criterion ID	Verbatim Wording	JESIP Principle(s) Mapped	Classification Rule
	ANT:9	Anticipation of wider incident implications – cover moves, road closures, early joint media strategy	Co-ordination; Shared situational awareness	Proxy
	DM:9	Multi-agency decision-making consideration in line with local protocols	Co-ordination; Communication	Explicit
	PLAN:2	Development of strategies that are aligned to objectives, and a joint multi-agency working strategy	Co-ordination; Communication	Explicit
	PLAN:5	Request of appropriate resources incl. local, regional, national, international arrangements	Co-ordination	Explicit
	PLAN:7	Planning of actions with internal FRS support functions and other responding/support agencies	Co-ordination	Proxy
	PLAN:8	Development & implementation of risk control/contingency measures and utilisation of safe systems of work	Joint understanding of risk	Proxy
	PLAN:9	Recording of essential information	Shared situational awareness	Proxy
	COMM:4	Use of safety briefings, where appropriate, with other agencies and FRS personnel	Communication	Explicit
	COMM:6	Effective Communication with local community – warn & inform	Communication; Shared situational awareness	Proxy
	COMM:7	Effective wider incident media management	Communication; Co-ordination	Proxy
	COMM:8	Communication of the incident situation to other responders via fire control using agreed formats	Communication; Shared situational awareness	Explicit
	CMD:8	Establishment & maintenance of the liaison with other agencies and consideration of JESIP principles	All five principles	Explicit
	CMD:9	Ensure systems for ongoing safety, sustainability, and welfare (and incident recovery)	Co-ordination; Review	Proxy
	REV:4	Review effectiveness of multi-agency actions	Co-ordination; Review	Explicit
	REV:5	Review of the effectiveness of current strategy and tactics	Co-ordination	Proxy
	REV:7	Review of incident information to assess effectiveness & sustainability of resources & capabilities	Shared situational awareness	Proxy
	REV:8	Evaluation of effectiveness of decisions & operations (independently and multi-agency)	Co-ordination; Joint understanding of risk	Explicit

Appendix B: Full descriptive statistics for all JESIP and non-JESIP behavioural markers.

	mean	std	25%	50%	75%	IQR
q1_1	3.29719	0.594081	3	3	4	1
q1_2	3.236061	0.558532	3	3	3	0
q1_3	3.235158	0.585863	3	3	3	0
q1_4	3.222952	0.574381	3	3	3	0
q1_5	3.212745	0.598595	3	3	3	0
q1_6	3.161603	0.517579	3	3	3	0
q1_7	3.153568	0.522174	3	3	3	0
q1_8	3.218145	0.552487	3	3	3	0
q1_9	3.116225	0.529588	3	3	3	0
AverageScoreForInformation	3.191178	0.428477	3	3	3.3	0.3
q2_1	3.238411	0.552385	3	3	3	0
q2_2	3.198914	0.544636	3	3	3	0
q2_3	3.228099	0.57373	3	3	3	0
q2_4	3.198458	0.56347	3	3	3	0
q2_5	3.235451	0.564794	3	3	3	0
q2_6	3.174653	0.568167	3	3	3	0
q2_7	3.173574	0.510704	3	3	3	0
q2_8	3.210136	0.570877	3	3	3	0
q2_9	3.162254	0.535299	3	3	3	0
AverageScoreForUnderstanding	3.188324	0.435954	3	3	3.3	0.3
q3_1	3.215575	0.558788	3	3	3	0
q3_2	3.216934	0.558251	3	3	3	0
q3_3	3.189025	0.541999	3	3	3	0
q3_4	3.192837	0.543874	3	3	3	0
q3_5	3.17829	0.548953	3	3	3	0
q3_6	3.19166	0.55511	3	3	3	0
q3_7	3.171589	0.521484	3	3	3	0
q3_8	3.160795	0.542177	3	3	3	0
q3_9	3.135617	0.524098	3	3	3	0
AverageScoreForAnticipation	3.170278	0.435609	3	3	3.2	0.2
q4_1	3.188602	0.548855	3	3	3	0
q4_2	3.22026	0.590855	3	3	3	0
q4_3	3.221196	0.546818	3	3	3	0
q4_4	3.237167	0.582533	3	3	3	0
q4_5	3.246861	0.568452	3	3	3	0
q4_6	3.208119	0.537052	3	3	3	0
q4_7	3.188153	0.521612	3	3	3	0
q4_8	3.17218	0.500151	3	3	3	0
q4_9	3.163332	0.518961	3	3	3	0
AverageScoreForDecisionMaking	3.192011	0.446192	3	3	3.3	0.3
q5_1	3.268037	0.574027	3	3	3	0
q5_2	3.211216	0.541498	3	3	3	0
q5_3	3.224546	0.559261	3	3	3	0

q5_4	3.227416	0.565868	3	3	3	0
q5_5	3.177575	0.550014	3	3	3	0
q5_6	3.181836	0.550602	3	3	3	0
q5_7	3.170619	0.506757	3	3	3	0
q5_8	3.173579	0.521739	3	3	3	0
q5_9	3.139256	0.553881	3	3	3	0
AverageScoreForPlan	3.183835	0.437994	3	3	3.2	0.2
q6_1	3.231605	0.558505	3	3	3	0
q6_2	3.24508	0.583094	3	3	3	0
q6_3	3.198295	0.578031	3	3	3	0
q6_4	3.164628	0.541556	3	3	3	0
q6_5	3.185518	0.54882	3	3	3	0
q6_6	3.161343	0.51	3	3	3	0
q6_7	3.189675	0.541412	3	3	3	0
q6_8	3.178784	0.567826	3	3	3	0
q6_9	3.181419	0.519973	3	3	3	0
AverageScoreForCommunication	3.179442	0.439056	3	3	3.2	0.2
q7_1	3.194782	0.541431	3	3	3	0
q7_2	3.18691	0.541442	3	3	3	0
q7_3	3.225685	0.57439	3	3	3	0
q7_4	3.215276	0.546861	3	3	3	0
q7_5	3.184113	0.566341	3	3	3	0
q7_6	3.19082	0.54285	3	3	3	0
q7_7	3.246967	0.58347	3	3	3	0
q7_8	3.140041	0.49638	3	3	3	0
q7_9	3.143192	0.48956	3	3	3	0
AverageScoreForCommand	3.178729	0.434861	3	3	3.2	0.2
q8_1	3.1876	0.53264	3	3	3	0
q8_2	3.196409	0.533455	3	3	3	0
q8_3	3.217267	0.548253	3	3	3	0
q8_4	3.182519	0.511756	3	3	3	0
q8_5	3.186656	0.515756	3	3	3	0
q8_6	3.152365	0.474465	3	3	3	0
q8_7	3.161701	0.503434	3	3	3	0
q8_8	3.159917	0.478485	3	3	3	0
q8_9	3.151064	0.466519	3	3	3	0
AverageScoreForReview	3.165939	0.417122	3	3	3.2	0.2

S.T.A.R.A. (Simple Triage Rapid Aid): A new protocol

Alexandre Sérgio de Oliveira Angelin PhD SMS-UNICAMP/SP®
Ednei Fernando dos Santos PhD UNICSUL/SP
Tiago Regis Franco de Almeida PhD UNIESP/SP
Marcelo Donizeti Silva Postdoctoral EERP-USP/SP

School of Medical Sciences
University of Campinas (Unicamp)
São Paulo- Brazil
Email alexandreangelin@yahoo.com.br / a271295@dac.unicamp.br

https://doi.org/10.61618/EFYD7229

Abstract

Introduction: Drowning fatalities pose a global public health challenge, particularly in Mass Casualty Incidents in Aquatic Environments (MCI-AqE). Ineffective screening victim triage is critical when there are many victims and resources are overwhelmed, and lack of planning.
Objective: This article introduces a Protocol for Simple Triage and Rapid Aid (S.T.A.R.A.). Aim is propose a theoretical triage model for aquatic environments, integrating variables like buoyancy, response time, and victim grouping. Methods: S.T.A.R.A.-Protocol was developed using a qualitative, descriptive method based on expert consensus and illustrative scenarios. Its formulation involved rapid scene assessment, victim observation, strategic floating equipment deployment, and prioritization of response time focusing on consciousness and aquatic mobility. The protocol uses five color-coded triage categories and a clear decision-making flow. Results: The simulation of S.T.A.R.A. demonstrates enhanced rapid in MCI-AqE. improved operational efficiency, structured resource allocation, and timely interventions. A protocol visual and color-coded system facilitates decision-making, optimizing rescue efforts. Conclusion: S.T.A.R.A.-Protocol offers a standardized, practical, and efficient framework for triage in aquatic mass casualty incidents in aquatic environments. It test is expected to improve rescue operations, reduce fatalities, and optimize resource allocation, enhancing preparedness and response capabilities. Beyond operational benefits, S.T.A.R.A. also supports rescuer mental health. By providing a clear, structured decision-making framework in high-stress MCI-AqE scenarios, it reduces uncertainty in cases for first responders, then better cognitive load. This systematic approach can mitigate the psychological impact of chaotic events that contribute to PTSD and other mental diseases in rescue personnel.

KEY WORDS: *Firefighters, Rescue, Triage, MCI, Aquatic environment.*

Introduction

Mass Casualty Incidents (MCI) or drownings involving Mass Casualty Incidents in Aquatic Environments (MCI-AqE) represent a significant challenge for rescue services worldwide. The unpredictability of catastrophic and fatal drowning MCI-AqE has drawn attention in several nations. The rapid and accurate triage is essential to prioritize care and optimize resources and many models were development and validation for MCI (Culley & Effken, 2010). In incidents, approaches and perspectives in the face of the challenges presented are emphasized by International Life Saving Federation (ILS, 2024), and in Brazil by Brazilian Water Rescue Association (SOBRASA, 2024). Despite the existence of established protocols, Bazyar, Farrokhi and Khankeh (2019) explain some types in around at world, and Culley, et al. (2014) explore five methods promise such as START by (Benson, Koenig and Schultz, 1996). However, when it comes to MCI-AqE, highlights the limited number of models tailored to water-based rescues. Recent studies by Barton, Morgan and Tipton (2024) explain this models. Others, Rescue (USAR) System (Caroline, 1992); Jump START (Romig, 2002); MCI Triage Triage (Lerner, Schwartz, & Coule , 2008); SALT (Triage (Lerner, 2008); Aircroff Chasher Into Water (Hickey, 2012); Shipwireks (Hansen, 2012); beach-related (Matthews, Andonaco & Adans, 2014). Effective response to MCI-AqE requires more than simply adapting land-based protocols. It demands a focus on frontline rescuers (Bierens, Knape & Gelissen, 2002; Quan et al., 2016), careful assessment of drowning victims, particularly those being swept away (Sempsrott et al., 2019), strategic deployment of flotation resources (Schimidt et al., 2016), and prioritization of response times (Papa, Hoelle & Idris, 2005). However, institutional preparedness often lags behind these needs. Catastrophic events can expose systemic failures and highlight the urgent need for innovative strategies that support rescuers and decision-makers in aquatic environments (Pascual-Gómez, 2011; Lopes, 2019). Catastrophes that devastate communities and modify the thinking of public institutions reflect a regression to a stage that can be characterized as a systemic failure in the response to a critical event. Case studies in Bangladesh (Rahman et al., 2009), Germany (Reijnen et al., 2018), South Africa (Saunders, Sewduth & Naidoo, 2018), and France (Castle et al., 2020) are examples. The resource allocation in chaotic aquatic environments, reducing cognitive load and mitigating post-traumatic stress among responders, and contributing to evidence-based practices, are base to offering a standardized of theory-driven triage model. Then, Study proposes as the S.T.A.R.A. (Simple Triage And Rapid Aid) protocol are a new conceptual model for triage in MCI-AqE. This reveals the level of institutional preparedness for critical events (Harichandan, 2023). Cultural changes are valid when they affect transformative epistemic structures, both of knowledge and praxis, as demonstrated by the panorama of tragedies around the world. Protocol aims to improve rescue operations and reduce fatalities by providing a structured decision-making framework, and seeks to standardize conduct and expand the operational capacity of rescuers in high-risk water rescue contexts. The proposal is contributed to the standardization of a protocol in MCI-AqE.

Literature Review

In the last ten years, few publications appear on work involving MCI-AqE. Discussions refer to MCI or suggest for new perspectives on aquatic safety. Many works focusing on attitudes towards drowning victims (Peden, Franklin, & Leggat, 2016) or even examining decision-making in aquatic disaster situations (Barton, Morgan & Tipton, 2024). The Terminology is a factor by analysis in these contexts.

Define for Multiple Casualty Incidents in Aquatic Environments (MCI-AqE).

Currently, there are multiple terminologies and methods for MCI Triage (Lerner, Schwartz, & Coule , 2008). New terms for casualty were developed as a management tool for emergency response. The aim is support decision makers in assessing casualty that needs based on limited information and finite resources, and minimizing disruption (Mills, et al., 2006). As example, a promissory management tools for emergency response that aimed at the Urban Aid and Rescue (USAR) System (Caroline, 1992). The author includes management tools for emergency response and many terminologies. The World Health Organization (WHO) defines MCI as an event that generates more casualties than available first responders, thus compromising the response capacity of agents (WHO, 2007). The Ministry of Health in Brazil, through the Mobile Emergency Care Service (SAMU), conceptualizes MCI as an incident involving a number greater than or equal to five victims (Brasil, 2016). And the Civil Defense of the State of São Paulo, objective of our study, defines it as *"sudden events that produce a number of victims that leads to an imbalance between available medical resources and needs, where maintaining an adequate standard of care is only possible through the mobilization of local resources*" (São Paulo, 2012, p. 4). In the context of MCI, Timbie et al. (2013) associate the response to such an incident with a service protocol that, given the limitation of resources and the slowness of actions, restricts the operational capacity of rescuers. When adding MCI with AqE, particularly in cases of drowning (Golden, Tipton , and Scott, 1997), water rescue (Szpilman et al., 2012), or near-drowning (Szpilman, 2000; Stephenson and Byard , 2023), the subject remains understudied in academia. Mass Casualty Incidents include shipwrecks (Hansen, Jepsen & Hermansen, 2012), aircraft crashes into water (Hickey, 2012) and beach-related incidents (Matthews, Andronaco & Adams, 2014). On populated beaches, factors such as shoreline length, accessibility, type of public or frequency, and user demographics should be considered to assess the probabilities of incidents (Pellicioni, 2014) as these factors may culminate in multiple drowning events. Notably, the WHO defines drowning as the process of experiencing respiratory impairment by submersion/immersion in liquid (Szpilman et al., 2012, citando OMS, 2002). Emergency response to these chaotic incidents can become dynamic and complex, requiring rescue services to demonstrate a high level of planning, resource coordination, and professional preparedness (Short & Hogan, 1994; Scott, 2007). Research indicates that the chances of survival decrease rapidly after submersion. According to the United States Lifesaving Association (USLA), field data from professional beach lifeguards suggests that a critical two-minute response window significantly increases successful rescue outcomes (USLA, 2022). Findings from the state of São Paulo emphasize a distance of 50 meters as the distance for a rescuer to reach the victim within the 3-minute intervention limit, considered the ideal rapid response time (Pellicioni, 2014). It is important to highlight the different levels of risk, primarily from the swimmer's history, behavioral patterns, operational efficiency of rescue teams, and environmental conditions. Furthermore, most critically, submersion time remains a critical factor in the survival of a drowning victim (Suominen , et al., 2002). Considering hypoxemia and respiratory failure caused by submersion in liquid, rapid intervention becomes crucial for the victim's survival. In most cases, cardiac arrest occurs within minutes, and hypothermia reduces cerebral metabolic activity due to decreased oxygen consumption (Brooks, 2001). Therefore, an immediate and effective response is vital, with cardiopulmonary resuscitation (CPR) maneuvers proving effective in combating hypoxemia. Others details, most drownings occur in warmer aquatic environments. Particularly at beaches, and represent the largest number of high-risk incidents for both rescue and drowning. Given the importance of drowning prevention, the establishment of optimized techniques and the standardization of knowledge dissemination have led to reflection on the development of a rapid triage protocol for incidents with multiple casualties in water, specifically in populated environments. Rio Grande do Sul, in Brazil, innovative research involve alternative methods of classifying aquatic victims (Oliveira et al., 2017; Szpilman, 2019). This method does not emphasize the standardization of triage, reduced response time and ease of application through a specific method. Then, enter as challenge for rescues. Considering the importance of drowning prevention, the establishment of best practices, standardized techniques and the dissemination of knowledge have led to the development of rapid triage protocols for multiple casualty incidents in aquatic environments. At moment, the START by (Benson, Koenig and Schultz, 1996) continue being the method propose.

Simple Triage and Rapid Treatment (START) Approach Protocol

To respond quickly the incidents involving multiple victims, one of the first protocols created was START by (Benson, Koenig and Schultz, 1996). The method is triage criteria applied for the purpose of saving lives on three main criteria: breathing, perfusion (circulation) and mental status (Benson, Koenig and Schultz, 1996). It represents the main triage protocol adopted by Fire Departments worldwide, at example Rescue Aquatic Guide of São Paulo State, Brazil (SÃO PAULO, 2006). The START was developed in the 1980s in California, USA, and considered a disaster triage method designed to quickly classify. Considered simple, fast, and systematic because it does not prioritize medical diagnosis but rather the rapid triage of victims at the hot scene. It uses the mnemonic "30-2-can do", which assesses “30”; respiratory rate in one minute “2”, peripheral perfusion quality in two seconds; and "Can do", the victim's level of consciousness (Bazyar, Farrokhi, & Khankeh, 2019).

Casualties are classified using a four-color coding system: Red, seriously injured casualties with a high probability of survival if treated immediately; Yellow, moderately injured casualties who can wait up to an hour for treatment; Green, occasionally "walking casualties" with minor injuries, able to ambulate and assist in emergency operations; and Black or Gray for dead or casualties in cardiopulmonary arrest with no viable chance of immediate recovery (Kahn, Schultz, & Anderson, 2009). Although the accuracy of the START method may be debatable, it demonstrates reasonable accuracy and is comparable to other triage models. It currently serves as a standard disaster triage tool (Franc et al., 2022). The effectiveness of the system stems from its incorporation of the intuitive mechanisms inherent in traditional victim assessment methods.

Proposed Method

In Brazil, considered necessary to develop a model capable of assessing the victim's condition even from a distance - whether on land or distant aquatic environments. This innovation allows rescuers to quickly triage victims according to the severity of the injury, thus optimizing the allocation of resources during MCI. The propose was to refining the steps by Oliveira et al. (2017), and then create the STARA protocol (Simple Triage And Rapid Aid) based on the fundamental principles of the START model status (Benson, Koenig and Schultz, 1996). Experts from Firefighters High School (FHS), Fire Department of the military in São Paulo State, Brazil and São Paulo Fire Department Commission were consulted. Considered a national reference in aquatic rescue in Brazil. Specialists of School of Nursing of São Paulo University (EN-USP), of Science Medical School of Unicamp (SMS-Unicamp) and Health Sciences from Cruzeiro do Sul University (UNICSUL/SP) were consulted yet. The contribution validates the conceptual consistency of the proposed method. It aims to provide efficient emergency management in the low time, allowing rescuers to save lives by prioritizing victims who would benefit most from rapid intervention Timbie et al. (2013). It must be emphasized that the STARA protocol is still at an idealized stage. To date, it represents a conceptual and theoretical construction and has not yet been formally validated. Practical simulations with firefighter school trainees have been carried out, which provide initial feasibility insights but do not replace real controlled operations.

Model Comparatives

In comparison with existing triage systems, among the most used, START (Benson, Koenig & Schultz, 1996) is the most recognized internationally, and prioritizes rapid categorization through basic vital signs assessment, focused on land scenarios. STARA adopts the simplicity of START but adapts it to victims located at a distance, where direct measurement of vital signs may not be immediately possible in water. Regarding flexibility and lifesaving interventions, Lerner et al. (2008) proposed the SALT (Sort, Assess, Lifesaving interventions, Treatment/Transport). It has already been used in comparisons with other systems, and Integrates lifesaving interventions during triage. STARA incorporates the structured decision logic of SALT but adjusts it for contexts where immediate interventions are not always feasible. The JumpSTART, created by Romig (2002) as a pediatric adaptation of START, was introduces rescue breaths before the final decision but otherwise follows the START logic of breathing, perfusion, and mental status. STARA not focus on age, instead adds other variables such as buoyancy, immersion time, and rescue logistics, focusing specifically on immersion-related factors. The SMART follows the same sequence as JUMP START. Created by Mawji et al. (2022) it introduces a more systematic use of structured flowcharts and the development of a rapid triage algorithm for use in low- and middle-income countries. STARA does not focus on income, but has adapted flowcharts for criteria decisions.

The SMART follows the same sequence as JUMP START. Created by Mawji et al. (2022) it introduces a more systematic use of structured flowcharts. Development of a rapid pediatric triage algorithm for use in low- and middle-income countries. The protocol SIEVE/SORT described by Bazyar (2019), applies rapid initial screening (SIEVE) followed by detailed classification (SORT). Similarly, STARA follows a two-phase logic: a preparation phase that considers future variables, and an aquatic operation phase that emphasizes pre-rescue assessment, followed by structured categorization after removal from the water. The protocol SIEVE/SORT described by Bazyar (2019), applies rapid initial screening (SIEVE) followed by detailed classification (SORT). Similarly, STARA follows a two-phase logic: a preparation phase that considers future variables, and an aquatic operation phase that emphasizes pre-rescue assessment, followed by structured categorization in water and after removal from the water. There are other traditional methods that show limitations. Continued with these as a basis for STARA, which was designed specifically to address operational gaps in aquatic environments, does not replace existing protocols but complements them by addressing a specific gap. Access, measurement, and intervention are delayed by the nature of the incident.

Most systems are validated through simulations rather than scientific data. For this reason, international cooperation is essential to achieve more organized responses to mass casualties (Wang et al., 2022). **Table 1** summarizes the context of application, aquatic limitations, and characteristics of each protocol.

Table 1. Comparative of Triage Protocols

Reference	*Protocol*	*Application*	*Main Features*	*Limitations in Aquatic Environments*
	STARA (Simple Triage and Rapid Aid)	Aquatic environments, MCI with drowning risk	Two stages (preparatory/observation - structured classification); flowchart with contextual variables	Still conceptual; no empirical validation, only simulations with trainees
Benson, Koenig & Schultz (1996)	START (Simple Triage and Rapid Treatment)	Land-based, general adult population	Rapid assessment (respiratory rate, perfusion, mental status)	Requires proximity to victim for vital signs; not feasible at distance or in water
Romig (2002)	JumpSTART	Land-based, pediatric focus	Modified criteria for children's physiology (e.g., apnea management)	Not designed for aquatic rescue; relies on direct contact with victim
Lerner et al. (2010)	SALT (Sort, Assess, Lifesaving interventions, Treatment/Transport)	Mass-casualty, flexible across scenarios	Integrates lifesaving interventions during triage	Immediate interventions often not possible in aquatic settings
Mawji et al. (2022)	SMART	Major incidents, systematic	Uses tags, structured flowcharts, color-coded categorization	Tagging impractical in aquatic incidents
Jafar Bazyar (2019) *	SIEVE/SORT	Pre-hospital, UK standard	Sieve -rapid primary triage; Sort - detailed secondary triage	Requires initial proximity and vital sign measurement

* Is widely referenced in literature on emergencies and mass triage (e.g., in Systematic Reviews and Comparative Studies).

A New protocol

The proposed method was developed based on a national and international literature review and analysis of existing protocols. Were observed Studies focusing on drowning incidents, disaster response, and mass triage, at example, protocol START (Benson, Koenig & Schultz, 1996), and aquatic rescue protocols, especially the SOBRASA model (Oliveira et al., 2017). By a qualitative and descriptive approach, the conceptual development of a new triage protocol applied for MCI-AqE. This explained that the authors' practical experience in aquatic rescue contexts caused the principal protocol START to become outdated, as Wang et al. (2022) describe, as the construction was only in an environment. The theoretical assessment and practical plausibility, conducted through empirical experiences of the authors and alignment with institutional references, WHO, FEMA, SOBRASA, led to the creation of the method to improve operational efficiency. For a creation, a critical analysis of operational gaps in existing triage systems was carried out, considering contextual variables of aquatic environments, such as distance, buoyancy, response time and victim behavior. Finally, a conceptual modeling of the STARA protocol, structured in two stage, preparatory stage and application stage. As a decision logic by flowchart, moving on to operational applicability in the field, classification criteria, and categorization by cores was elaborated. **Figure** 1 explain the design conceptual model.

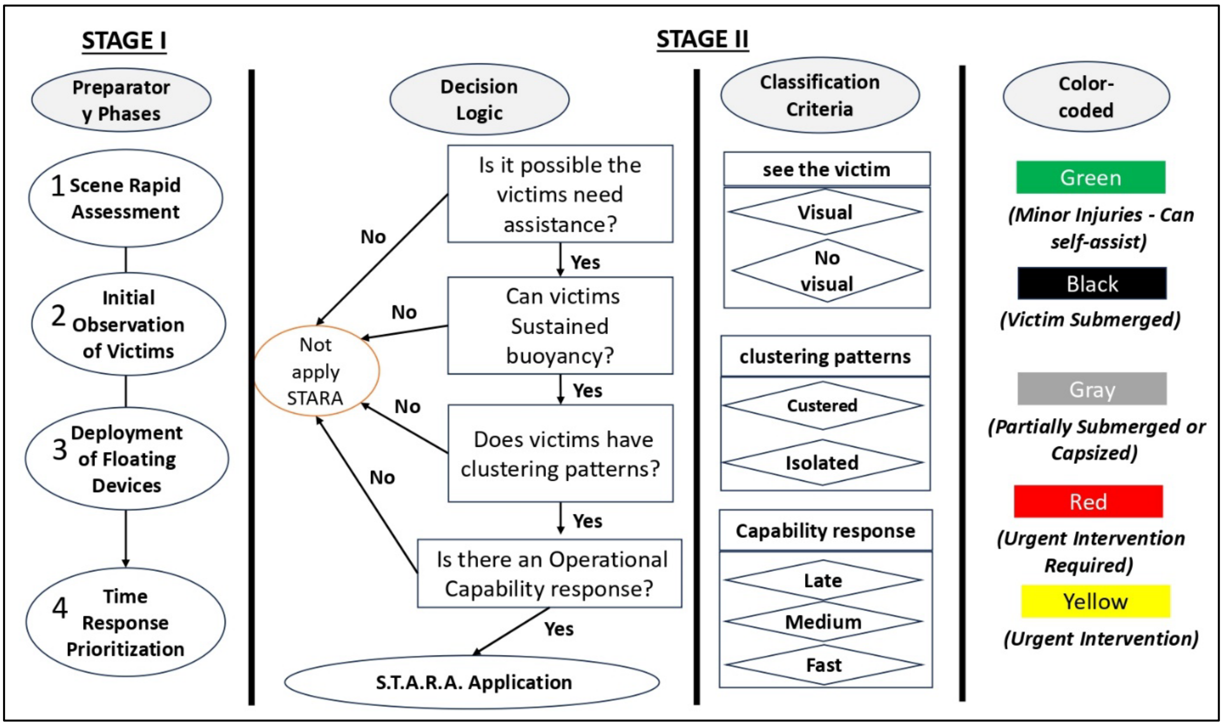

Figure 1: Desing conceptual modeling of the STARA protocol

The main limitation identified concerns the absence of real empirical trials. The data does not replace controlled operational validation, then it is a conceptual proposal. Details, the STARA stages are more looking if composed of two main divisions: before and after the call for help. Regardless of the actions, Stage I is executed with the preparation of the place by a professional and the execution of the routine. It includes making a note of the observed scene, looking for potential peoples, prioritizing resources, and facilitating time responsiveness. For Stage II, after the call for help, occur application of the protocol occurs, and Stages I and II together. To learn to understand the stages after a distress call.

STARA Stage I – Preparatory Phase (before)

STARA consists of four preparator y phases. These are: (1) Scene Rapid Assessment (SRA); (2) Initial Observation of Victim (IOV); (3) Deployment of Flotation Device (DFD); and (4) Time Response

Prioritization (TRP). **Figure 1** illustrate the phases. The first step consists of a quick assessment of the scene. It involves a perception of the profile of victims. Figure 2 exemplifying the sea, beach, and the profile of swimmers and bathers. The professional must have active access to the operational area, a Strategic Positioning, an Environmental Analysis, and a Zone Delimitation, in which the rescuer defines a Responsibility Limit Zone (ZLR). The ZLR for MVI is an imaginary line that divides hazards into primary and secondary or surf zone (ZI), as Castelle et al. (2018) define. Primary requires rapid action by the responder; secondary requires slightly more elaborate planning, with other forces or recurs assisting. Another case, for example, a shipwreck, ZLR is within a radius of 15 to 45 meters from the vessel. In the next phase, rescuers begin to outline coordinated actions to act in prevention or rescue processes. The professional has active access to the operational area, with knowledge of the currents, channels, and profile of swimmers. In summary, the stage consists of the rescuer's perception of potential victims when observing from a distance. The next step consists of initially observing victims, if in groups or concentrated, and rescuers' perception of potential victims when observing from a distance. Try to understand the groups of people and whether they are grouped in pairs, threes, or fours. Note how many groups there are. This occurs through the analysis of the behavioral profile of potential victims, by characteristics or competence of their swimming and patterns of engagement in the water. In this situation, a mechanism of continuous assessment by visual tracking and activation of progressive response is necessary, since the assessment of the threat of drowning is in real time. The third phase involves Deployment of floating devices. Flotation Equipment Strategy for emergency response, where the provision of floating equipment is coordinated to reach more victims with less equipment. It must be prepositioned at strategic points and such flotation devices focused on response time, incident probability, and/or inserted in service units. Equipaments must be allocated in a way that serves the majority of those involved. This ensures that resources are more easily accessible when intervention is necessary. (Timbie, et al., 2013). For example, on a boat, there are life jackets for the people on board. Beaches may have lifeguards carrying out prevention with boats. In an emergency, all equipment must be taken into consideration weighing the degree of risk according to the number of victims. Technological resources can also be incorporated. For example, coast of São Paulo, the Command and Special Operations Center (CCSO), a 24-hour surveillance and monitoring center, and through it, it is possible to use video surveillance of the beach to respond to the call in the event of an emergency drowning or eliminate the need for support (Botelho, 2019). The fourth stage consists of Time Response Prioritization. This phase focuses on the response to victims, through three main dimensions: distance of the rescuer for possible intervention or positioning myself, awareness of the buoyancy of the observed groups, and degree of consciousness and unconsciousness of the groups or actual buoyancy. The three dimensions affect the ability to save lives by the time responders.

STARA II Stages – Status Assessment Phases (after)

This stage involves the application of the protocol itself. The concept is widely applied in disasters related to victims of various types of drowning. The responder must have started Stage I, and make a logical drawing, asking: Is it possible the victims need assistance? If yes, Can victims sustained bourancy? If yes, Do victims have a cluster pattern? If yes, Is there an Operational capability response?

If yes, STARA Apply. If no in anywhere, not STARA Apply. After, the rescuer observes the Classification criteria. See the victims? visual or no-visual; See Clustering patters? Clustered or isolated; Capability response? Late, medium ans fast. Then it is possible to Structural System in the mind during the movement to the victims. Within, argue the Need for Assistance (NA). NA allows the rescuer to continually monitor victim groups, recognizing that they remain in a state of struggle, dealing with the lingering physiological and psychological effects of the incident, oscillating between resilience and vulnerability. When NA is "no" and "Isolated", the rescuer imagines that the victim can save himself.

After, Aquatic competence (AC) refers a situation where the rescuers have visual or no-visual of victims, based on the isolation pattern or clustering Pattern (CP). If there is no visual, there is no reason to search for the victim. If there is visual and the victim is alone or in a group, this provides the simultaneous understanding of next idea, which is Sustained Buoyancy (SB). SB refers to the active floating capacity of the victim(s) and a certain level of swimming ability. The actual buoyancy that the victim(s) have keeps them floating in the water and serves to alleviate immediate suffering until help arrives. In other words, it is a question of whether or not the victim has aquatic competence to survive.

After, the Grouping Pattern (GP) refers to the dynamics or spatial organization of victims during an Adverse Drowning Event (AEA). This pattern is related to the emergent self-organization among drowning victims, highlighting social and individual needs in the context of providing mutual assistance - "clustered" or "isolated". To this end, throughout the process, rescuers continuously monitor victims to recognize whether they remain in a state of struggle and whether victims should be grouped. It is worth noting that the group's survival capacity creates collective resilience through shared situational understanding and temporary bonds (Erfurth et. al., 2021).

Finally, Response Capacity (RC), which refers to the operational factor of the rescuer, not in relation to the victim, but rather to the rescuer. Through it, a multidimensional construct is constituted that encompasses the operational skills of the rescuer during aquatic emergencies. The professional or team must be fully aware of their own ability, efficiency, and readiness to respond. It has four components: a) response capacity; b) readiness; c) speed; and d) training.

Response Capacity - involves the ability to organize, plan and execute the necessary and available resources at the work site (Liu, Z., & Tian, H., 2025). *Readiness* - refers organization's ability to keep professionals positioned or ready to act in emergency situations. It stands out in situations involving professionals at strategic points or situations involving boats in matrix support (White, 2012). *Speed* - is related to the real time of your response and should be understood as between 2 to 4 minutes for fast time, 4 to 6 minutes for medium time, and more than 6 minutes for late time (Pellicioni, 2014). *Training* - refers to the professional's performance training in the face of the impositions that the organization orders or to which professionals are subject. It is up to the professional to align their own well-being, their physical, mental and intellectual health to respond to emergencies in different critical scenarios as a first responder (Rooke, A., & de Terte, I., 2024). It is clear that all actions are closely related to technical preparation, and based on the interval between the rescuer's visualization and arrival at the victim's position. And, when aiming at performance, the color-coded triage system was created in order to bring efficiency to the emergency response, with the aim of ensuring that victims receive fast, efficient and humane care (FEMA, 2021).

Color-coded sorting systems

As previously reported, current ground casualty triage methods employ color-coding as a mechanism to facilitate impact for the incident (Benson, Koenig & Schultz, 1996). Although Color is personal and subjective, the use of color as a visual guide determines the appropriate course of action and increases the efficiency and clarity of the response. It is used for situational and behavioral analysis. FEMA highlights that color-based systems not only speed up the processing for rescue victims but also maximize survival rates (FEMA, 2021) and can be used in aquatic environments. Is a universal communication that allows the severity of each victim's condition to be recognized and the rescuer to quickly assess the scene. The process occurs dynamically and responders are guided by decision-making (visual processing of ≤ 3 seconds) and by optimizing the priority sequencing of the entire response process (De Medeiros Dantas, et. al., 2022). In STARA, the victim's status is organized into five priority categories, represented by the colors Red, Yellow, Green, Gray and Black. Based in START by (Benson, Koenig & Schultz, 1996), category Green (Priority One) is a stable condition with a high probability of survival. Victims demonstrate excellent swimming ability. They are close to safety and show no signs of distress. They are able to escape the danger zone without immediate assistance. Others who require rapid support are instinctive fighters. For example, an individual or group of experienced amateur swimmers in areas far from the surf zone who are successfully escaping the area. Category Yellow (Priority Two) is a moderate condition requiring assistance. Victims have noticeable difficulty swimming, remain conscious, and maintain basic water skills. They may be isolated, but mobile victims require immediate assistance. If not rescued, they should await rescue, especially in isolated cases. In groups, mutual support may allow continued floating. Evidence of a level of cooperation allows delayed intervention in some cases. For example, one person or a group in deep water helping each other stay afloat. Category Red (Priority Three) is the critical state with risk of drowning. Although victims are grouped together, they are actively drowning, with severe difficulty in staying afloat and ineffective mutual assistance. Urgent, rapid/medium intervention by a rescuer is required. Alternate between submersion and attempts to keep the head above water. Immediate rescue is required due to compromised buoyancy and uncoordinated movements. There is clear motor disorganization among them. The risk is compounded by the fact that one victim may involuntarily pull others underwater. For example, victims clinging to each other while repeatedly submerging.

As previously reported, Color action and increases the efficiency and clarity of the response, and current ground casualty triage methods employ color-coding as a mechanism to facilitate impact for the incident (Caroline, 1992); (Benson, Koenig & Schultz, 1996); (Romig, 2002); (Lerner, Schwartz, & Coule , 2008); (Hickey, 2012); (Hansen, 2012); (Matthews, Andonaco & Adans, 2014).

The Gray Category (Reduced Priority) is the Critical State with life-threatening conditions, since victims are found face down, without motor activity and with their face submerged. It represents an extreme emergency with imminent risk to life, requiring immediate intervention and advanced life support. Entangled victims may be considered. For example, victims lying face down in the water. There may be respiratory arrest in these cases, without motor activity. The Black Category (Priority Zero) is when the victim is not visually located or initially sighted, with visual loss after contact with water, or victims lost sight of. These are cases resulting from complete submersion or based on indirect reports. This classification applies to situations with a total absence of visual reference at the time of rescue (Lerner, et al., 2011). For example, sudden submersion witnessed without return to the surface, and the last known position becomes the epicenter of the search. By combining the color-coded sorting systems and Status Assessment Phases, we can create a flowchart for a quick explanation of the method. **Figure 2** illustrates the application of the phases of assessing the condition of the victims and the application of the color system by STARA Protocol Structural.

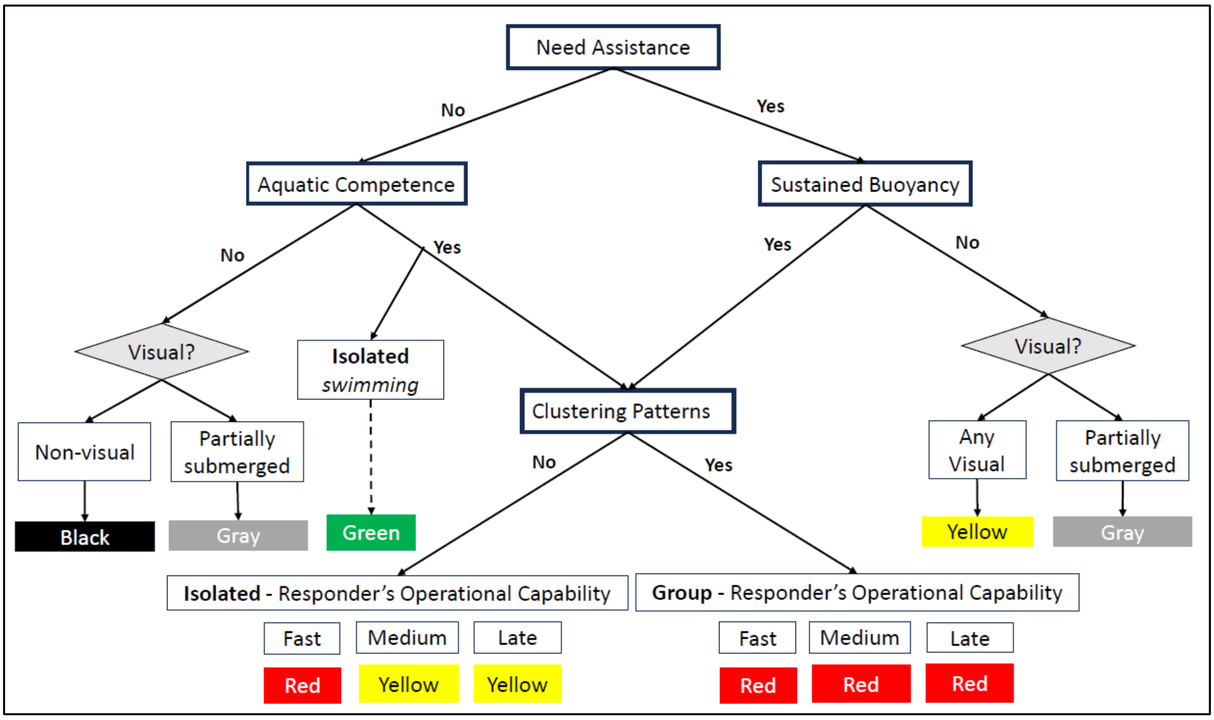

Figure 2: STARA Protocol Structural

Considerations about the Protocol

STARA relies on the integration of theory and practice, mainly for its adaptability to dynamic variations, taken from other methods. Applicability of the STARA method must be through an intuitive decision-making flowchart designed for high-pressure and complex situations. By understanding and memorizing, and training, the logic of the flowchart guides responders step-by-step, for a better rescue. As described in **Figure 1**, the construction of the STARA protocol resulted in a systematization divided into four preparatory phases preceding the method. Rescuer identifies the degree of urgency based on visual and assessments, and uses STARA as tool for fast and efficient decision-making in complex situations, which enhances the emergency operations. As mentioned, its application is based on prioritizing care according to the severity of the situation (group or isolated) and the probability of survival of the victims (Sustained Buoyancy). It uses the final categorization of the color system, such as green, yellow, red, gray, and black. Potential for diverse applications, which can be on beaches, dams, shipwrecks, and even flooded urban settings. Something that differs from protocols already consolidated in the academic and professional environment is incorporates innovative concepts, such as the need for intervention, Aquatic competence, Sustained Buoyancy, and clustering Pattern analysis assessment. Highlighting that the model does not substitute the START, whose application in aquatic environments still presents limitations (Benson, Koenig and Schultz, 1996).

Practical Application

Two simulated in MCI-AqE scenarios are presented. These examples are based on situations frequently faced by rescue teams and demonstrate how the application of the protocol can guide quick and effective decisions.

Scenario 1 – Urban beach on the weekend

Situation: Sunday, 2pm, urban beach with a large flow of bathers. A group of 7 teenagers is swept away by a rip current. The rescuer identifies from a distance three victims actively struggling, close to each other, and another two trying to swim far away from each other out of the danger zone.

Preparatory acts: Stage I had been prepared because the Lifeguard was already close to the scene.

STARA application: the rescuer approaches the scene, and the victims are observed. First, triaged with a quick response time (1 to 2 min); Equipment is deployed directly with the entry of three floats into the area of greatest risk. There are three victims with loss-visible of buoyancy and grouped together (Red Category), and two others are swimming approximately 10 meters away (Yellow Category). Target red victims. Second, providing immediate support for the yellow victims who can await secondary rescue or wait for the priority of response.

Decision principle: The decision was based on the grouping and the observed buoyancy rate to rescue of victims.

Scenario 2 – Urban flooding with multiple victims

Situation: Monday, 14pm, A Flood suddenly hits a riverside neighborhood as least 12 people seen on rooftops, others on improvised buoys.

Preparatory acts: Operations Center relaying information and moving teams.

STARA application: Phase I in action. Team observes the scene from a high point. The professionals end up knowing the location and identifying victims. Get driving floating equipment to the scene. Organizing themselves in priority of action according to the ZLR. First, the condition out of the Responsibility Limit Zone (ZLR) imposed by the team commander serves as a base for the rescue of victims. Observation of the victim group. Three adults centered on a buoy with their feet in the water (Red Category). Conscious adults and children organized in groups on the roof (Red Category). The other five are near the security area in a backwater area (Yellow Category). Two children visibly capsized, partially submerged (Gray Category). Second, there must be a request for immediate support for every victims. Third, A main responder observes the victims who can be removed. Another focuses on the equipment to be used, such as rafts, floats, and life jackets for everyone. Team A focuses on deploying equipment whose essence is on the victims with the buoy and those on the roof. The response priority, in which Red Victims are triaged with a quick response time (2 to 4 min) because the risk of being swept away is high. Team B Yellow victims. Removal with use of the equipment to accommodate everyone. Finally, after everyone has been removed, the focus is on capsized victims or victims missing in the water, with emergency diving teams.

Decision principle: The team is directed to the victims categorized as red. The decision was based on the grouping and risk of death due to the current carrying the buoy to another point that is distant. And the second team prioritizes the rescue of the largest number of victims from the roof. After, the victims were categorized as yellow. People groups in a backwater area. Subsequently, after removing all the living victims, start being guided by radio and a response strategy focused on the search for victims capsized or missing in waters, categorized as gray or black.

Expected results

The practical application of the STARA protocol was illustrated through two simulated scenarios, one involving an incident on an urban beach and the other in a flooded area.

In both cases, the rescuers approached the scene and began the triage of the victims by STARA. In case two there was a preparation of Phase I on-site.

In the first scenario, adolescents are at risk of drowning due to rip currents. The buoyancy assessment allowed the identification of victims actively struggling and at imminent risk (rapid response and red category as priority). others as the presence of isolated victims, conscious, and visual were categorized as yellow priority.

n the second scenario, an urban flood that exposes multiple victims on rooftops and floating objects. The triage observed the formation of spontaneous groups. The first was on a buoy, potentially exhausted, at risk of being swept away by the current. These were categorized as immediate priority, red. The second, those kept on the rooftop and isolated, at risk of the building collapsing or worsening the flood situation, were categorized as immediate priority (red). Others were categorized yellow.

Finally, partially submerged children were categorized as gray. Theses require action after the rescue of the living or the use of an emergency diving team (additional team activated via radio).

In both contexts, the responders applied core tasks of the STARA protocol. Assessing observable buoyancy (yes/no), identifying clusters (yes/no), and determining response capacity based on estimated time to intervention (fast, medium, or late).

Then, prioritization decisions can be guided by a logical flowchart coded by clusters, which allows victims to be quickly categorized according to risk. Action based on this protocol demonstrated the potential to increase the efficiency of visual tracking and direct available resources to victims, minimizing the risk of death by drowning.

Conclusion

Currently, operations involving MCI-AqE face four main challenges, namely accurate assessment of victims, professional performance in the field, availability of flotation equipment, and prioritization of response time. The gap in standardized protocols in the context of incidents involving multiple victims makes it difficult to implement effective and practical actions at the aquatic scene. Therefore, STARA emerges as a new perspective on triage in MCI-AqE responding to a specific gap in life-saving practices. By considering factors such as buoyancy, equipment, and response time, STARA enhances the decision-making capacity of first responders in critical situations, where seconds define individuals' lives. Rapid categorization of victims using color-coded systems and objective urgency criteria, with a focus on increasing the effectiveness of rescue efforts while mitigating the psychological impact on victims, families, and rescuers, has a particular impact on reducing the risk of emotional distress. Include the emotional distress and this reduction of cognitive load and uncertainty can mitigate the psychological impact of chaotic events. This help visual behavioral analysis of rescuers supporting post-traumatic stress disorder (PTSD) prevention among rescue personnel and promoting a healthier, more resilient emergency workforce.

Future Research

Encourage researchers, educational institutions, and rescue agencies to test, adapt, and improve STARA protocol. In order to ensure that faster and more efficient rescues can result from the collaboration between theory and practice, so that lives can be saved , we suggest that future research explore the integration of STARA with surveillance and tracking technologies, as well as its impact on response times and victim survival rates. Furthermore, there is a need for empirical validation of the model. Therefore, new on-site applications are recommended in order to evaluate the adaptability of the protocol. Variables such as distances, environmental conditions and resource allocation can be observed. Future research and effective interventions, aligned with international experience, promote significant advances, then the STARA can strengthen as a significant protocol in the management of aquatic emergency response systems. The theoretical construction and conceptual modeling presented in this article provide a solid basis. The STARA proposal goes beyond its potential application in the aquatic environment, making it an invitation to create responsive, intuitive and adaptable systems to the new emergencies of the 21st century.

Limitations

The S.T.A.R.A. protocol remains at a conceptual and theoretical stage. Although its design was informed by international literature and practical experience of rescue professionals, it has not yet undergone formal empirical validation or testing in real-world scenarios with actual victims. The current evidence derives mainly from simulations with firefighter school trainees and controlled training exercises, which provide only preliminary feasibility insights. In the comparative analysis with established protocols such as START (Benson, Koenig & Schultz, 1996) and the Brazilian SOBRASA models (Oliveira et al., 2017), the START remains the international reference for land-based triage but requires direct access to victims and vital signs, which is not feasible in water environments, and the protocol by Oliveira et al. (2017) provides a national perspective on aquatic rescue but does not emphasize standardized triage or reduced response time. The STARA aims to complement these approaches by adapting their principles to aquatic scenarios, filling an identified operational gap. However, until STARA is validated with empirical data, its comparative advantages remain hypothetical. Field studies in different aquatic settings and collaboration with international agencies should be necessary to confirm its reliability, safety, and adaptability under real conditions. This lack of operational validation introduces potential biases in the assessment of its effectiveness and applicability. Any conclusions about its impact on response times, victim outcomes, and rescuer safety should be interpreted with caution. This protocol should be viewed as an initial proposal, a structured model, rather than as a validated or definitive standard for aquatic mass casualty triage.

ACKNOWLEDGMENTS

The authors read and approved the final manuscript and declared no conflicts of interest. This project did not receive grant funding. The lead author is available for further clarification on the protocol.

ABOUT THE AUTHOR

Alexandre Angelin is a PhD candidate in Public Health at the University of Campinas (UNICAMP), Brazil. His research focuses on occupational mental health, post-traumatic stress disorder (PTSD), and suicide prevention, specifically among emergency responders such as firefighters and police officers. He is currently developing and implementing a pilot project on MCI-AqE at São Paulo Firefighters College, Brazil, combining theoretical models with practical field strategies to improve emergency response outcomes. For correspondence, please contact the Department of Epidemiology, School of Medical Sciences, University of Campinas, São Paulo, Brazil; SP or mail: Phone: +55 16 9963 0303. Email address: alexandreangelin@yahoo.com.br; alexandreangelin@policiamilitar.sp.gov.br a271295@dac.unicamp.br .

References

Barton, C. A., Morgan, P., & Tipton, M. J. (2024). Development of a novel 'In-Water Mass Casualty Triage Tool'. *BMJ Military Health*. [Internet]. Available from: https://pure.port.ac.uk/ws/portalfiles/portal/98213292/Development_of_a_novel_In-Water.pdf

Bazyar, J., Farrokhi, M., & Khankeh, H. (2019). Triage systems in mass casualty incidents and disasters: A review study with a worldwide approach. *Open Access Macedonian Journal of Medical Sciences, 7*(3), 482–494. https://doi.org/10.3889/oamjms.2019.119

Benson, M., Koenig, K. L., & Schultz, C. H. (1996). Disaster triage: START, then SAVE—A new dynamic triage method for victims of a catastrophic earthquake. *Prehospital and* Disaster Medicine, 11 (2), 117–124.

Bierens, J. J., Knape, J. T., & Gelissen, H. P. (2002). Drowning. *Current Opinion in Intensive Care, 8*(6), 578–586.

Brazil. Ministry of Health. Secretariat of Health Care. (2016). *Intervention protocols for SAMU 192* (2nd ed.). Brasília: Ministry of Health. Retired: [Internet]. [cited 2024-09-14]. Available from: www.samu.gov.br

Brazilian Water Rescue Society (SOBRASA). (2024). *Drowning - Epidemiological Bulletin in Brazil and National Strategic Plan*. Recuperado em 20 de outubro de 2024, de https://sobrasa.org/dados-sobre-afogamento/

Caroline, N. L. (1992). Disaster response: Principles of preparation and coordination. *JAMA, 268*(16), 2306–2307.

Castelle, B., Brander, R., Tellier, E., Simonnet, B., Scott, T., McCarroll, J. & Lechevrel, P. (2018). Surfing hazards and injuries on beaches in southwestern France. *Natural Hazards, 93*, 1317–1335. https://doi.org/10.1007/s11069-018-3354-4

Castelle, B., Scott, T., Brander, R. W., & McCaroll, R. J. (2020). Wave and tidal controls on rip current activity and drowning incidents in southwestern France. *Journal of Coastal Research, 95*(SI), 769–774. https://doi.org/10.2112/SI95-150.1

Cone, D. C., Serra, J., Burns, K., MacMillan, D. S., Kurland, L., & Van Gelder, C. (2009). Teste piloto do Sistema SALT de Triagem de Vítimas em Massa. Atendimento Pré-Hospitalar de Emergência, 13(5), 536–540. https://doi.org/10.1080/10903120802706252

Culley, J. M., & Effken, J. A. (2010). Development and validation of a mass casualty conceptual model. *Journal of Nursing Scholarship, 42*(1), 66–75. https://doi.org/10.1111/j.1547-5069.2009.01320.x

Culley, J. M., Svendsen, E., Craig, J., & Tavakoli, A. (2014). A validation study of 5 triage systems using data from the 2005 Graniteville, South Carolina, chlorine spill. *Journal of Emergency Nursing, 40*(5), 453–460. https://doi.org/10.1016/j.jen.2014.04.020

De Medeiros Dantas, Í. J., Batista, FEA, Solino, LJS, Freire, AG, do Nascimento, MN, & Júnior, GS (2022). A dimensão psicológica das cores: uma revisão sistemática da literatura sobre psicologia das cores. *Pesquisa, Sociedade e Desenvolvimento*, *11* (5), e34111528027-e34111528027.

Erfurth, LM, Hernandez Bark, AS, Molenaar, C., Aydin, AL, & van Dick, R. (2021). If worse comes to worst, my neighbors come first": social identity as a collective resilience factor in areas threatened by sea floods. ***SN Social Sciences*** *1* (11), 272. DOI: https://doi.org/10.1007/s43545-021-00284-6

Federal Emergency Management Agency (FEMA). (2024). *Emergency Management and Training Manuals*. [cited 2024 Oct 29]. Retrieved from https://www.fema.gov

Franc, J. M., Kirkland, S. W., Wisnesky, U. D., Campbell, S., & Rowe, B. H. (2022). METASTART: A systematic review and meta-analysis of the diagnostic accuracy of the Simple Triage and

Rapid Treatment (START) algorithm for disaster triage. *Prehospital and Disaster Medicine, 37*(1), 106–116. https://doi.org/10.1017/S1049023X2100131X

Hansen, H. L., Jepsen, J. R., & Hermansen, K. (2012). Factors influencing survival in the event of shipwreck and other maritime disasters in the Danish merchant fleet since 1970. *Safety Science, 50*(7), 1589–1593. https://doi.org/10.1016/j.ssci.2012.03.016

Harichandan, P. (2023). Institutional disaster preparedness: A case study of Odisha. In *International Handbook of Disaster Research* (pp. 83–94). Singapore: Springer Nature Singapore.

Hickey, J. J. P. (2012). *A theoretical system safety analysis of US Coast Guard aviation mishap involving CG-6505* (Doctoral dissertation). Massachusetts Institute of Technology.

International Life Saving Federation (ILS). (2007). *ILS Drowning Report*. [cited 2024 Set 20]. Retrieved from https://www.ilsf.org/drowning-prevention/conference/wcdp2023/

Kahn, C. A., Schultz, C. H., Miller, K. T., & Anderson, C. L. (2009). Does START triage work? An outcomes assessment after a disaster. *Annals of Emergency Medicine, 54*(3), 424–430. https://doi.org/10.1016/j.annemergmed.2008.12.035

Lerner, E. B., Cone, D. C., Weinstein, E. S., Schwartz, R. B., Coule, P. L., Cronin, M., ... & Hunt, R. C. (2011). Mass casualty triage: an evaluation of the science and refinement of a national guideline. *Disaster medicine and public health preparedness*, *5*(2), 129-137. https://doi.org/10.1001/dmp.2011.39

Lerner, E. B., Schwartz, R. B., Coule, P. L., Weinstein, E. S., Cone, D. C., Cronin, M., … Hunt, R. C. (2008). Mass casualty triage: An evaluation of the data and development of a proposed national guideline. Disaster Medicine and Public Health Preparedness, 2(S1), S25–S34. https://doi.org/10.1097/DMP.0b013e318182194e

Lin, Y.-K., et al. (2022). Simple triage and rapid treatment protocol for emergency department mass casualty incident victim triage. *The American Journal of Emergency Medicine, 53*, 99–103. https://doi.org/10.1016/j.ajem.2021.12.037

Liu, Z., & Tian, H. (2025). The formation mechanism of emergency response capacity in major emergencies at the grass-roots level. *Eurasia Journal of Science and Technology.* Volume 7, Issue 4, Pp 62-71, 2025

Matthews, B., Andronaco, R., & Adams, A. (2014). Beach warning signs: Do they work? Safety Science, 62, 312–318. https://doi.org/10.1016/j.ssci.2013.09.003

Mawji, A., Li, E., Dunsmuir, D., Komugisha, C., Novakowski, S. K., Wiens, M. O., … Ansermino, J. M. (2022). Smart triage: Development of a rapid pediatric triage algorithm for use in low-and-middle income countries. Frontiers in Pediatrics, 10, 976870. https://doi.org/10.3389/fped.2022.976870

Papa, L., Hoelle, R., & Idris, A. (2005). Systematic review of definitions for drowning incidents. Resuscitation, 65(3), 255–264. https://doi.org/10.1016/j.resuscitation.2004.11.030

Quan, L., Bierens, J. J., Lis, R., Rowhani-Rahbar, A., Morley, P., & Perkins, G. D. (2016). Predicting drowning outcome at the scene: A systematic review and meta-analyses. Resuscitation, 104, 63–75. https://doi.org/10.1016/j.resuscitation.2016.04.006

Rahman, A., Mashreky, S. R., Chowdhury, S. M., Giashuddin, M. S., Uhaa, I. J., Shafinaz, S., et al. (2009). Analysis of the fatal childhood drowning situation in Bangladesh: Exploring prevention measures for low-income countries. Injury Prevention, 15(2), 75–79. [cited 2024 Nov 11]. Retrieved from https://injuryprevention.bmj.com/content/15/2/75.short

Reijnen, G., Buster, M., & Vos, J. (2018). Epidemiological aspects of drowning and nonfatal drowning in the waters of Amsterdam. Journal of Forensic and Legal Medicine, 58, 78–81. https://doi.org/10.1016/j.jflm.2018.04.014

Romig, L. E. (2002). Pediatric triage: A system to JumpSTART your triage of young patients at MCIs. JEMS: Journal of Emergency Medical Services, 27(7), 52–60.

Rooke, A., & de Terte, I. (2024). The working life of a surf lifesaver: The traumatic experiences and consequences of New Zealand surf lifeguards.*Traumatology, 30*(1), 58–68. https://doi.org/10.1037/trm0000243

Sanddal, T. L., Loyacono, T., & Sanddal, N. D. (2004). Effect of JumpSTART training on immediate and short-term pediatric triage performance. Pediatric Emergency Care, 20(11), 749–753. https://doi.org/10.1097/01.pec.0000144912.49827.6f

São Paulo. Military Police of the State of São Paulo. (2006). *Aquatic rescue guide*. Fire Department of the State of São Paulo. São Paulo, Brazil.

Saunders, C. J., Sewduth, D., & Naidoo, N. (2018). Keeping our heads above water: A systematic review of fatal drownings in South Africa. South African Medical Journal, 108(1), 61–68. [cited 2024 Dec 07]. Retrieved from https://www.ajol.info/index.php/samj/article/view/166359

Schmidt, A. C., Sempsrott, J. R., Hawkins, S. C., Arastu, A. S., Cushing, T. A., & Auerbach, P. S. (2016). Wilderness Medical Society practice guidelines for the prevention and treatment of drowning. *Wilderness and Environmental Medicine, 27*(2), 236–251. https://doi.org/10.1016/j.wem.2015.12.019

Scott, T., et al. (2007). Beach rescue statistics and their relation to nearshore morphology and hazards: A case study for southwest England. *Journal of Coastal Research, 50*(sp1), 1–6. https://doi.org/10.2112/JCR-SI50-001.1

Short, A. D., & Hogan, C. L. (1994). Rip currents and beach hazards: Their impact on public safety and implications for coastal management. *Journal of Coastal Research*, 197–209. [cited 2024 Oct 29]. Retrieved from https://www.jstor.org/stable/25735599

Suominen, P., et al. (2002). Impact of age, submersion time and water temperature on outcome in near-drowning. *Resuscitation, 52*(3), 247254. https://doi.org/10.1016/S0300-9572(01)00478-6

Szpilman, D. (2000). Drowning. *Brazilian Journal of Sports Medicine, 6*, 131–144. https://doi.org/10.1590/S1517-86922000000400005

Timbie, J. W., Ringel, J. S., Fox, D. S., Pillemer, F., Waxman, D. A., Moore, M., et al. (2013). Systematic review of strategies for managing and allocating scarce resources during mass casualty events. *Annals of Emergency Medicine, 61*(6), 677–689.e101. https://doi.org/10.1016/j.annemergmed.2013.02.005

United States Lifesaving Association (USLA). (2022). *Public Education*. [cited 2024 Jun 22].. Retrieved from http://www.usla.org

Wang, J., Li, X., Zhang, Y., Chen, L., & Liu, M. (2022). The usage of triage systems in mass casualty incident of developed countries. Open Journal of Emergency Medicine, 10(2), 124–137. https://doi.org/10.4236/ojem.2022.102011

White, Jill E. (2012). *Starguard:* Best practices for lifeguards. Human Kinetics.

World Health Organization (WHO). (2007). *Mass casualty management systems: Strategies and guidelines for building health sector capacity*. Genebra: World Health Organization. [cited 2024 Dec 11]. Retrieved from http://www.who.int

Letter To the Editor: Public Drone Use and Its Impact on Search and Rescue and Wildfire Operations

Toby Meredith PhD, Elizabeth Cuevas PhD

Email: jsar@journalofsar.com

https://doi.org/10.61618/KVWX6292

This letter was co-developed by both authors. Following the passing of Toby Meredith during the drafting phase, the final manuscript was completed by his co-author, Elizabeth Cuevas, consistent with the direction they established together.

Public use of unmanned aircraft or drones continues to affect emergency response operations in ways that responders, aviation teams, and agencies cannot ignore. Recreational drone ownership has expanded rapidly, and more individuals are flying these devices near active disaster scenes. Although many operators believe they are assisting, uncoordinated drone flights introduce risks that slow operations and place both responders and survivors in danger. As researchers and practitioners working in disaster operations, we have observed a significant rise in uncoordinated drone activity at active response sites. This letter aims to highlight that trend and call on response agencies, regulators, drone manufacturers, and the public to address this issue directly.

The July 2025 floods in the Texas Hill Country demonstrate how quickly these hazards can develop. As helicopter crews conducted hoist operations and reconnaissance in unstable conditions, unauthorized drones entered the airspace. Several near misses were reported, and one rescue helicopter in Kerr County was struck by a drone and forced to land, removing a vital aircraft from service during an active rescue cycle (KSAT News, 2025). Responders noted that unauthorized drones complicated flight paths and reduced available decision time during aerial search operations (DroneLife, 2025; Stokel-Walker, 2025).

Similar interference has been documented internationally. In January 2025 a privately operated drone collided with a Super Scooper aircraft working an active wildfire in California, causing damage significant enough to remove the aircraft from service during suppression efforts (Los Angeles Times, 2025). Unmanned aircraft sightings reported by aircrews indicate a steady increase in unsafe operations, with pilots reporting evasive maneuvers in nearly three percent of encounters in 2025 (Wallace, 2025).

The risks of drone collisions with helicopters or fixed wing aircraft are well established. Federal Aviation Administration impact testing shows that even lightweight drones can damage rotors, engines, or windshields upon impact (Federal Aviation Administration, 2017). When pilots see or suspect a drone in their airspace, they must slow, alter, or temporarily suspend flight operations until the area is confirmed safe, which introduces complex operational risks. These delays also reduce the speed and effectiveness of rescues and wildfire suppression.

Although many recreational operators intend to help, uncoordinated drone flights do not support responders. Incident commanders cannot verify or integrate imagery or data collected by personal drones, and such information may conflict with operational formats or create liability concerns. During Hurricane Harvey, unauthorized drones initially interfered with air operations, prompting the Texas Military Department to warn publicly that “civilian drones pose EXTREME risks to our rescue pilots and crews in high need areas” (IoT World Today, 2017). After this warning, several volunteer drone groups worked with agencies to coordinate flight patterns, ground aircraft upon request, and share imagery only through official channels. This shift improved safety and demonstrated that civilian groups can contribute meaningfully when they operate under unified direction rather than independently. The DroneUp partnerships used during Harvey further illustrate how civilian operators, when organized and aligned with official command, can support search efforts without adding risk (AirSight, 2017). That experience underscores the difference between organized support and self-directed flight.

Legal restrictions prohibiting drone operation near disaster scenes exist in the United States, the United Kingdom, Canada, and Australia. Enforcement remains challenging. Responders cannot divert personnel to locate drone operators during an active emergency, and counter-drone technologies are not universally available or authorized for local agencies.

Despite these concerns, drones have meaningful value when deployed within coordinated response systems. Agencies in the United Kingdom, Canada, Australia, and the United States use drones to map fire behavior, document search areas, assess structural conditions, and deliver real-time situational awareness. Research from Sweden shows that automated external defibrillator-equipped drones arrived before ambulances in a majority of trials and reduced time to first shock by nearly two minutes, demonstrating the potential of well-integrated drone systems to save lives (Karolinska Institute, 2020).

Emergency response agencies should strengthen public education that emphasizes how unauthorized drone flights restrict aviation safety. The West Midlands Fire Service has set a strong example with its direct messaging urging the public to keep personal drones away from emergency scenes (West Midlands Fire Service, 2023). Regulators and manufacturers should expand geofencing and develop automatic restrictions that activate during declared emergencies. Agencies should also create accredited pathways that allow trained civilian pilots to support operations safely rather than through uncoordinated flights.

Uncoordinated drone flights place responders and communities at unnecessary risk. As severe flooding, drought conditions, and wildfire activity increase across multiple regions, safe and predictable airspace will remain essential to effective emergency aviation. Keeping uncoordinated, personal drones grounded during active incidents is a necessary and achievable step toward protecting both responders and the people they are working to reach.

References

AirSight. (2017). *Drones, the law, and Hurricane Harvey: A case study*. https://www.airsight.com/blog/case-study-drones-the-law-and-hurricane-harvey/

DroneLife. (2025, July 21). *Drones join the front lines in deadly Texas Hill Country flood response*. https://dronelife.com/2025/07/21/drones-join-the-front-lines-in-deadly-texas-hill-country-flood-response/

Federal Aviation Administration. (2017). *UAS airborne collision study*. U.S. Department of Transportation.

IoT World Today. (2017). *Private drones ordered to stay away from Harvey rescue efforts*. https://www.iotworldtoday.com/transportation-logistics/private-drones-ordered-to-stay-away-from-harvey-rescue-efforts

Karolinska Institute. (2020). *Automated external defibrillator-equipped drones for facilitating early defibrillation in out-of-hospital cardiac arrest* [Doctoral dissertation]. Karolinska Institutet. https://openarchive.ki.se/articles/thesis/Automated_external_defibrillator-equipped_drones_for_facilitating_early_defibrillation_in_out-of-hospital_cardiac_arrest/26900668

KSAT News. (2025, July 7). *Drone strikes rescue helicopter over Hill Country flood zone*. https://www.ksat.com/news/local/2025/07/07/drone-strikes-rescue-helicopter-over-hill-country-flood-zone/

Los Angeles Times. (2025, January). *Drone collides with firefighting aircraft*.

National Interagency Fire Center. (2024). *Standards for fire and aviation operations (Red Book)*. U.S. Department of the Interior. https://www.nifc.gov/sites/default/files/redbook-files/RedBook_Final.pdf

Wallace, R. J. (2025, September 16). *Testimony before the United States House of Representatives Committee on the Judiciary, Subcommittee on Crime and Government Surveillance: Unmanned and unchecked, confronting the rising threat of malicious drone use in America.* Embry-Riddle Aeronautical University

Stokel-Walker, C. (2025, July 8). Civilian drones hinder flood disaster relief in Texas. *Fast Company*. https://www.fastcompany.com/

Wallace, R. J. (2025). Why the Super Scooper wildfire aircraft drone collision should be a wakeup call to the drone industry. *DroneLife*. https://dronelife.com/2025/02/11/drone-incidents-involving-aircraft-should-be-industry-wakeup-call/

West Midlands Fire Service. (2023). *Please keep your drones away from emergency incidents*. https://www.wmfs.net/news/please-keep-your-drones-away-from-emergency-incidents-plea-from-west-midlands-fire-service/