applied sciences

MDPI

Article

Extending Human–Machine Interaction Analysis from Autonomous Driving to Manned–Unmanned Vehicle Teaming: A Function-Specific Effectiveness Framework and the Partial-Autonomy Trap

Giwhyun Lee [1], HyeonJun Yun [1], Jin-woo We [2] and Hongsuk Park [3,*]

[1] Department of Industrial Engineering, Kumoh National Institute of Technology, Gumi 39177, Republic of Korea; giwhyunlee@kumoh.ac.kr (G.L.); hjun0828@kumoh.ac.kr (H.Y.)
[2] Center for Defense Human Resource Management, Korea Institute for Defense Analyses (KIDA), Seoul 02455, Republic of Korea; wi0823@kida.re.kr
[3] School of Advanced Industry Convergence, Kumoh National Institute of Technology, Gumi 39177, Republic of Korea
* Correspondence: avipak@kumoh.ac.kr

Abstract

Human–machine interaction (HMI) has become a central issue in autonomous driving because partial automation can degrade, rather than improve, human performance during takeover, handover, and out-of-the-loop transitions. Similar interaction risks are emerging in manned–unmanned vehicle teaming (MUM-T), where operators must supervise multiple autonomous or remotely controlled assets under higher mission complexity and safety-critical constraints. However, existing effectiveness analyses of unmanned and MUM-T systems often treat the level of autonomy (LOA) as a fixed system attribute or assume the highest autonomy level, thereby obscuring the human–automation bottlenecks that arise during partial autonomy. This study proposes a function-specific HMI effectiveness framework in which autonomy is represented as a vector across surveillance, maneuver, fire or neutralization, command and control, and human–machine teaming functions. Measures of performance (MOPs) are modeled as conditional performances jointly shaped by function-specific LOA, operational environment, and intrinsic system capability, and are propagated through a five-layer LOA–MOP–MOE structure to a mission-level measure of effectiveness (MOE). The framework is demonstrated using a notional mine countermeasure scenario in which manned minehunters cooperate with unmanned underwater and surface vehicles. Four autonomy progression stages, from manned-centric operation to advanced cooperative autonomy, are evaluated for timely route opening. The case illustrates a non-monotonic partial-autonomy trap, or LOA-2 valley: remotely controlled unmanned assets may temporarily reduce mission effectiveness when teleoperation workload and HMI bottlenecks outweigh equipment gains, before cooperative and supervisory autonomy restore and exceed baseline performance. The contribution of this study lies not in the notional numerical results but in providing an explicit diagnostic structure for identifying where and why function-specific autonomy, control sharing, and HMI bottlenecks shape mission effectiveness. The framework thereby extends human–machine interaction analysis from autonomous driving to the broader, higher-risk setting of manned–unmanned vehicle teaming.

Academic Editor: Jun Chen

Received: 29 June 2026
Revised: 20 July 2026
Accepted: 22 July 2026
Published: 28 July 2026

Keywords: human–machine interaction (HMI); manned–unmanned teaming (MUM-T); level of autonomy (LOA); control sharing; out-of-the-loop performance; measure of effectiveness (MOE); measure of performance (MOP); mine countermeasures (MCMs); function-specific analysis; reliability, availability, maintainability, and cost (RAM-C)

1. Introduction

Human–machine interaction (HMI) has become a defining concern in the development of autonomous driving. As road vehicles progress along the SAE levels of driving automation, control is increasingly shared between the human driver and the automated system, and the quality of this interaction—rather than raw automation capability alone—often determines safety and performance. A recurring finding is that partial automation does not improve human performance monotonically: during takeover, handover, and out-of-the-loop (OOTL) transitions, drivers who have been disengaged from the control loop exhibit degraded situation awareness and slower, less stable responses when control is returned to them [1–3]. In particular, the relationship between automation exposure and takeover performance has been shown to be non-monotonic rather than linear [3], and the quality of control transitions depends on interaction factors such as operator workload, trust calibration, and the specific take-over parameters involved [1,2]. These results indicate that introducing automation can transiently degrade, rather than enhance, overall human–automation performance.

These interaction risks are not confined to civilian road vehicles. As autonomy extends from automated driving to defense applications, manned–unmanned teaming (MUM-T) systems—in which a human operator supervises and cooperates with one or more unmanned vehicles—face the same class of human–automation interaction challenges, but under higher mission complexity, safety-critical constraints, and adversarial conditions. In such systems, a partially autonomous or remotely controlled unmanned vehicle can shift workload onto the operator and introduce supervisory and coordination bottlenecks, so that the net effect of adopting automation on mission effectiveness is not necessarily positive. Yet, effectiveness analysis of unmanned and MUM-T systems has largely developed within a weapon-system effectiveness tradition that treats autonomy as a fixed precondition rather than as an interaction variable. Understanding how partial autonomy reshapes effectiveness therefore requires extending the HMI perspective from automated driving to manned–unmanned vehicle teaming—the aim of this study.

In the defense domain, MUM-T denotes an operational concept in which manned and unmanned platforms execute missions complementarily within a common operational environment. It has emerged as a key operational capability across diverse mission domains, including surveillance and reconnaissance, maneuver and breaching, fire support, security and protection, and mine countermeasures (MCMs) [4,5]. The effectiveness of such systems depends not only on platform performance but also on operational factors, such as the command and control (C2) mode, the division of roles between manned and unmanned components, human–machine teaming (HMT), and the cooperative architecture of multiple platforms [6]. Among these factors, the level of autonomy (LOA)—defined as the degree to which an unmanned system can perceive its environment, make judgments, and execute mission functions without continuous human intervention—is a key variable shaping the measures of effectiveness (MOEs) of unmanned and manned–unmanned systems [7].

A range of definitions and classification schemes for LOA have been proposed across various domains, including human factors engineering, automotive engineering, and military robotics. Representative examples include Sheridan's automation level model [8,9],

the SAE driving automation levels [10], and military-specific or multicriteria evaluation models, such as the Autonomy Levels for Unmanned Systems (ALFUS) framework [7,11], the North Atlantic Treaty Organization (NATO) Industrial Advisory Group (NIAG) framework for unmanned aerial vehicle (UAV) autonomy [12,13], and hierarchical models for unmanned platforms developed in subsequent literature [14–16]. These schemes have made important contributions to the conceptualization and systematic classification of autonomy by structuring it into discrete stages and explaining shifts in human–machine role allocation. However, they have generally interpreted autonomy in terms of technological maturity or the reduction of human intervention [17,18]. Consequently, they have shown limitations in explaining how changes in LOA alter mission-functional measures of performance (MOPs), such as those associated with surveillance and detection, maneuver, fire or neutralization, C2, and HMT and how these changes propagate to MOEs.

Furthermore, the US Department of Defense (DoD) Mission Engineering (ME) framework provides a Critical Operational Issue (COI)–MOE–MOP hierarchy for mission-based effectiveness analysis and serves as a common analytical framework for weapon-system effectiveness evaluation [19]. From an ME perspective, MOEs are mission-level indicators that assess whether "the system is doing the right things," whereas MOPs are function- and system-level performance indicators that assess whether "the system is operating properly" [19]. Several studies have explored the effectiveness of MUM-T and unmanned weapon systems using broader operational research frameworks, including agent-based modeling for amphibious operations, combat effectiveness analyses of unmanned surface vehicles, and the army weapon-system effectiveness analysis model (AWAM) for battalion-level Army TIGER analyses [20–22]. However, most existing effectiveness studies, including those grounded in the ME framework, have typically treated LOA as a scenario precondition or a fixed attribute. Consequently, explicit modeling of autonomy as a variable that influences MOPs has been limited [23,24].

Taken together, the limitations of prior work can be organized into three recurring problems. The first problem (P1) is the representation of autonomy as a single, system-wide scalar. The operator-involvement scales—Sheridan's ten-level model, the SAE driving-automation levels, and the hands-on-time frameworks of AIAA and AAD [4,8,10,16]—share this limitation directly, and the four-stage extension of Parasuraman et al. [9] distinguishes generic stages of human–automation interaction but does not decompose autonomy by mission function. Although the multicriteria frameworks—ALFUS, MAP, ACL, and the Pyramid and Cobweb models [7,11,14,15,25]—introduce multiple dimensions, these dimensions are generally not retained as mission-function-specific variables throughout an MOP–MOE effectiveness chain: some produce an aggregate score, whereas others preserve a multidimensional profile without specifying how the profile propagates to mission effectiveness. The second problem (P2) is that, in effectiveness evaluation, autonomy is typically omitted, fixed as a scenario condition, or implicitly treated as sufficiently mature for the evaluated endpoint, as in the reviewed effectiveness studies of amphibious, surface, land, and coastal-defense systems [20–24], which model equipment performance explicitly but leave autonomy outside the MOP structure. The third problem (P3) is that the human–machine interaction reshaped by autonomy—operator workload, supervisory control, and the division of labor between operator and system—is rarely represented explicitly as a performance-shaping variable in MOP–MOE structures. This gap spans both of the preceding groups: the classification schemes treat the interaction as a design characteristic rather than as a variable that shapes mission performance, and the reviewed effectiveness studies leave it unmodeled even where operator involvement is acknowledged as a limiting factor [20–24]. Yet, the literature on automated driving shows that precisely these interaction factors govern performance under partial automation [1–3].

The framework proposed in this study addresses each problem directly. P1 is addressed by defining autonomy as a function-specific vector across surveillance, maneuver, fire or neutralization, C2, and human–machine teaming functions, so that uneven maturation across functions remains visible. P2 is addressed by treating LOA as an explicit conditional variable that shapes MOPs jointly with the environment and intrinsic system capability, rather than as a scenario assumption. P3 is addressed by elevating human–machine teaming to a first-class function whose LOA-dependent coefficient propagates operator workload and supervisory constraints through the MOP structure to the mission-level MOE.

Against this background, this study proposes an analytical framework for the effectiveness evaluation of unmanned and MUM-T systems that is designed to address P1–P3. In this framework, LOA is interpreted not as a single system-wide level but as a function-specific autonomy vector and MOPs are defined as conditional performances that emerge from the joint effects of LOA, environmental conditions, and intrinsic system performance. The applicability of the proposed framework is illustrated through a minehunter-centered MCM case modeled using equation-based methods. In the case, a fully manned configuration serves as a baseline (A0) and three progressive MUM-T configurations A1–A3 represent the introduction of remotely controlled, semiautonomous, and autonomous unmanned assets, respectively.

This study does not aim to present a high fidelity quantitative model that predicts the actual combat effectiveness of a specific weapon system. Instead, it proposes a case-based analytical framework for identifying and interpreting the structural pathways through which autonomy influences mission effectiveness. In particular, by representing LOA as a function-specific vector and treating MOPs as conditional variables dependent on autonomy, the framework can reveal non-monotonic effectiveness pathways that are typically obscured under single system-wide LOA assumptions. These pathways include scenarios in which the partial introduction of remotely controlled unmanned assets may temporarily degrade rather than improve mission effectiveness.

The remainder of this paper is organized as follows. Section 2 presents a comparison and an analysis of the major LOA classification schemes proposed in industrial and military domains and summarizes their limitations. Section 3 presents a review of prior research on the ME-based composition of MOEs and MOPs and the incorporation of autonomy into MOE- and MOP-based frameworks. Section 4 proposes an LOA–MOP–MOE linkage framework based on a function-specific LOA vector and the conditional MOP concept. Section 5 presents (a) the application of the proposed framework to a minehunter-centered MCM case, (b) an analysis of how function-specific LOAs propagate through MOPs to mission-level effectiveness, and (c) a comparison with a conventional scalar-LOA analysis. Section 6 discusses the implications and advantages of the framework. Section 7 concludes the paper by summarizing the research findings, limitations, and directions for future research.

2. LOA Classification Schemes and Their Limitations

2.1. Autonomy as Function-Specific Human–Machine Interaction

The development of MUM-T is reshaping operational concepts and mission execution in the contemporary battlefield, with autonomy serving as a central driver of this transformation. Autonomy is not merely a degree of automation characterized by reduced human intervention; rather, it denotes the extent to which a system can perceive its situation, exercise judgment, and act independently within a given mission environment. It encompasses the degree of independence in a system's judgment and decision-making functions, the quality of human–machine collaboration, architectural efficiency, environmental

adaptability, and self-learning capacity. Accordingly, it constitutes a central factor in the operation of modern weapon systems and should be understood not merely as a technical attribute but as a structural factor shaping MOEs [7,19].

Underlying every level-of-autonomy scheme is a more basic question of human–machine interaction (HMI): which perception, situation-assessment, decision, and action responsibilities reside with the human operator, and which are delegated to the system. Raising the LOA does not simply remove the human; it redistributes roles, workload, and supervisory responsibility across these stages, changing the character of the interaction rather than eliminating it. Critically, autonomy need not advance uniformly across these stages or across the distinct functions a mission comprises: a system may be highly autonomous in perception yet remain operator-dependent in decision or action. Autonomy is therefore better understood as a function-specific property of human–machine interaction than as a single attribute of a platform.

Despite this conceptual centrality, definitions and classification criteria for LOA lack standardization across academic and technical domains. This fragmentation reflects multiple factors, such as cross-national variations in technological capability, differences in operational culture, and security-related constraints on technology sharing. Consequently, prior research has largely focused on developing autonomy metrics tailored to individual platforms, such as industrial automation systems, autonomous vehicles, UAVs, and unmanned ground vehicles (UGVs). However, it has been less effective in explaining the cooperative structures that govern MUM-T in multidomain and multi-echelon operational environments. In this context, existing autonomy classifications are useful for indicating technology maturity or function-level automation but share two common limitations: they represent human–machine interaction as a single, system-wide scalar rather than a function-specific condition, and they do not adequately explain the operational pathways through which autonomy affects MOE indicators, such as mission efficiency, accuracy, and reliability [18].

On this basis, this section presents a (a) review of the principal concepts and theoretical frameworks proposed for LOA classification and (b) comparative analysis of the representative schemes proposed in prior research. This review identifies structural commonalities and differences among these schemes and explores directions for developing LOA as an analytical framework linked to MOE-based effectiveness evaluation.

2.2. Comparative Analysis of LOA Classification Schemes

Research on LOA became more systematic with Sheridan's automation level theory in the 1980s. Since then, LOA research has expanded into autonomous driving, aerospace, and defense. Various classification schemes have also been proposed. However, these schemes differ markedly in terms of the criteria they use and their evaluation logic. In this section, we review principal LOA schemes by distinguishing industry-based classifications centered on human intervention from military-specific approaches. The latter are further divided into single-axis and multicriteria methods. Subsequently, we examine their structural commonalities and limitations [7,14].

2.2.1. Industry-Based LOA Classifications

The most widely cited classification scheme in industrial automation and control engineering is Sheridan's ten-level automation model proposed in 1992 [8]. This model arranges system autonomy along a continuum from Level 1 (fully manual) to Level 10 (fully autonomous) based on the degree of operator involvement. As shown in Table 1, these levels are defined by whether the human or the computer performs each step in a cycle comprising information acquisition, interpretation, alternative generation, and decision-making.

Level 1 corresponds to a fully manual state wherein all tasks are performed by the human operator. Levels 2–4 present decision alternatives—ranging from a full set of options to a single suggestion—but final approval remains with the operator. From Levels 5 to 7, the system executes actions under operator supervision, with prior approval, time-limited veto, or post-execution notification, respectively. At Levels 8 and 9, the system performs most decision-making and execution tasks autonomously, informing the operator only upon request or at its own discretion. Level 10 represents complete autonomy wherein the system performs all functions independently without human involvement.

Table 1. Sheridan's Ten Levels of Automation (adapted from Sheridan, 1992 [8]).

Level	Automation Level Criterion	Description
1	Manual	The computer provides no assistance; all decisions and actions are performed by the human operator.
2	Full set of alternatives	The computer provides a complete set of decision or action alternatives.
3	Restricted choice provided	The computer narrows the available alternatives to a limited set.
4	Single suggestion	The computer suggests a single alternative.
5	Execution upon approval	The computer executes the suggested action if approved by the operator.
6	Veto with time limit	The computer executes automatically unless vetoed by the operator before execution.
7	Automatic with notification	The computer executes automatically and then informs the operator.
8	Notification on request	The computer informs the operator only upon request.
9	Discretionary notification	The computer informs the operator only if it determines notification is necessary.
10	Fully autonomous	The computer performs all decisions and actions autonomously without human involvement.

As the first framework to classify autonomy according to the degree of operator involvement, Sheridan's ten-level model established the foundation for subsequent LOA research. The model was later adopted across industrial and engineering fields and strongly influenced later approaches to human–machine interaction. In the automotive sector, SAE International adopted this approach in its Levels of Driving Automation standard. However, because the SAE model was developed for civil road transportation systems, it is not directly applicable to MUM-T operations, which involve complex operational settings and coordination among multiple systems.

Subsequently, Parasuraman et al. extended Sheridan's model by dividing automation into four functional stages: information acquisition, information analysis, decision and action selection, and action implementation [9]. By distinguishing operator involvement across functions, this approach provided a more detailed view of human–machine interaction compared with earlier LOA models. However, it still interprets autonomy changes mainly in terms of technological advancement, thereby retaining the limitations of earlier LOA models.

Moray et al. applied the LOA concept to process control. They observed that as automation levels increase, direct operator involvement decreases, while the importance of supervisory control increases [17]. In other words, higher LOA does not simply reduce operator workload; it changes the operator's role and responsibility within the system.

Walker et al. analyzed LOA in the context of robot swarm operations and showed that a single LOA is insufficient. Instead, automation must vary across mission tasks and operator workload [6]. Their findings suggest that autonomy should not be viewed as a simple step-by-step progression but as a dynamic condition shaped by mission complexity and decision load.

Research based on Sheridan's model has substantially contributed to quantifying LOA according to operator involvement and to explaining changes in automation functions and operator roles. However, such approaches do not directly explain how changes in LOA affect MOEs. In particular, this line of research treats LOA as a stage of technological advancement or operator involvement rather than as an operational factor that shapes mission performance through MOPs.

2.2.2. Military-Specific LOA Approaches

While industry-based models define autonomy mainly in terms of operator involvement, military LOA research has increasingly focused on operational independence and mission execution capability. Military-specific approaches extend beyond the technical level of automation and examine how mission responsibility is shared between the human operator and the system. In particular, they place greater emphasis on role allocation and cooperation during mission execution. Military LOA approaches can generally be divided into two categories on the basis of methodology: (i) single-axis approaches, which classify autonomy using a single criterion such as operator involvement or control level, and (ii) multicriteria approaches, which evaluate autonomy by integrating factors such as mission complexity, environmental difficulty, and cooperative structure.

(1) Single-Axis Approaches: AIAA, AAD, UCAV, and FCS

The American Institute of Aeronautics and Astronautics (AIAA) and the United Kingdom's Aerospace, Aviation and Defense (AAD) Knowledge Transfer Network (KTN) each proposed six-level LOA frameworks based on pilot hands-on time [4,16]. Although these approaches differ in terminology and level composition, both assume that system autonomy increases as operator involvement decreases. The AIAA framework evaluates autonomy by quantifying pilot hands-on time, whereas the AAD framework places greater emphasis on the technical scope and functional capability of automation. Table 2 presents a comparison of the LOA criteria used by these two organizations. In both frameworks, higher levels correspond to reduced operator involvement and increased system autonomy. In the AIAA framework, LOA 0 corresponds to remote-controlled operation with 100% hands-on time, followed by simple automation, remote operation, highly automated or semiautonomous operation, full autonomy, and collaborative operation. The AAD framework follows a similar progression from manual control to collaborative operation.

Table 2. Comparison of Six Autonomy Levels: AIAA VSP HALE and AAD KTN (Young et al., 2005 [16]; AAD Knowledge Transfer Network, 2012 [4]).

LOA	AIAA VSP HALE	LOA	AAD KTN
0	Remote control (hands-on time 100%)	1	Manual control
1	Simple automation (hands-on time 80%)	2	Simple automation
2	Remote Operation (hands-on time 50%)	3	Remote operation
3	Highly automated or semiautonomous operation (hands-on time 20%)	4	Advanced automation or semiautonomous operation
4	Full autonomy (hands-on time below 5%)	5	Full autonomy
5	Collaborative operations	6	Collaborative operations

AIAA and AAD identify collaborative operations among multiple unmanned systems as the highest LOA. This reflects the growing importance of intersystem collaboration as a key dimension of autonomy in modern weapon system evaluation. From the perspective of MUM-T operations, autonomy must therefore be evaluated in terms of (a) operator involvement and (b) operational integration across multiple systems.

The US Defense Advanced Research Projects Agency (DARPA), through its Unmanned Combat Aerial Vehicle (UCAV) program, classified autonomy into four levels based on operator involvement [5]. Level 1 corresponds to manual operation, Level 2 to management by consent, Level 3 to management by exception, and Level 4 to fully autonomous operation. Each level is distinguished by the extent of human-in-the-loop control. The US Army's Future Combat Systems (FCS) program later extended the DARPA model into a ten-level LOA framework. Unlike earlier approaches, the FCS model considers not only operator involvement but also communication capability, cooperative capability, and autonomous mission execution. Consequently, it expands LOA from a simple automation scale into a mission-oriented evaluation framework. Table 3 presents a comparison of the LOA frameworks proposed in the UCAV and FCS programs.

Table 3. Comparison of LOA Frameworks in the UCAV and FCS Programs (adapted from National Research Council, 2005 [5]).

LOA	DARPA, US Air Force, and Boeing X-45 UCAV Programs	LOA	US Army FCS Program
1	Manual operation	1	Remote control
		2	Remote control with vehicle state knowledge
2	Management by consent	3	External preplanned mission
		4	Knowledge of local and planned path environments
		5	Hazard avoidance or negotiation
		6	Object detection, recognition, avoidance or negotiation
3	Management by exception	7	Fusion of local sensors and data
		8	Cooperative operations
		9	Collaborative operations
4	Full autonomy	10	Full autonomy

Note: The correspondence between UCAV and FCS levels shown here is the authors' approximation; NRC (2005) [5] presents the two scales separately.

As shown in Table 3, the DARPA program defines autonomy mainly in terms of operator involvement, whereas the FCS model incorporates mission functions such as situational awareness, hazard avoidance, and cooperative operations into a more detailed LOA framework. This shift suggests that military approaches to autonomy classification have moved beyond simply reducing operator involvement through technological advancement and have increasingly emphasized mission effectiveness and intersystem cooperation. However, these approaches still classify LOA along a technical progression without specifying the operational pathways through which increases in autonomy translate into changes in MOEs and MOPs.

In parallel with the aforementioned studies, several researchers proposed technical system capabilities, in addition to operator involvement, as key criteria for evaluating autonomy. For example, Sub-Group 75 of NATO's Industrial Advisory Group classified UAV autonomy into four levels on the basis of learning capacity and the observe–orient–decide–act (OODA) loop decision structure [12]. Meanwhile, Suresh and Ghose identified communication stability and information-sharing architecture as core elements of autonomy [13]. These technology-centered approaches addressed some limitations of earlier operator-centered models. However, autonomy in real operational environments cannot be adequately explained by any single criterion. Consequently, these approaches gradually

evolved into multicriteria frameworks that evaluate autonomy through the integration of multiple factors.

(2) Multicriteria Approaches: ALFUS, MAP, and Autonomous Control Level (ACL)

To address the limitations of single-axis frameworks, military LOA research introduced multicriteria approaches that evaluate autonomy using multiple dimensions. A representative example is the ALFUS framework developed by the US National Institute of Standards and Technology, which evaluates autonomy along three dimensions [7]: (i) mission complexity, (ii) human independence, and (iii) environmental complexity. Each dimension is defined on a quantitative scale. Figure 1 illustrates the three-dimensional structure of the ALFUS framework.

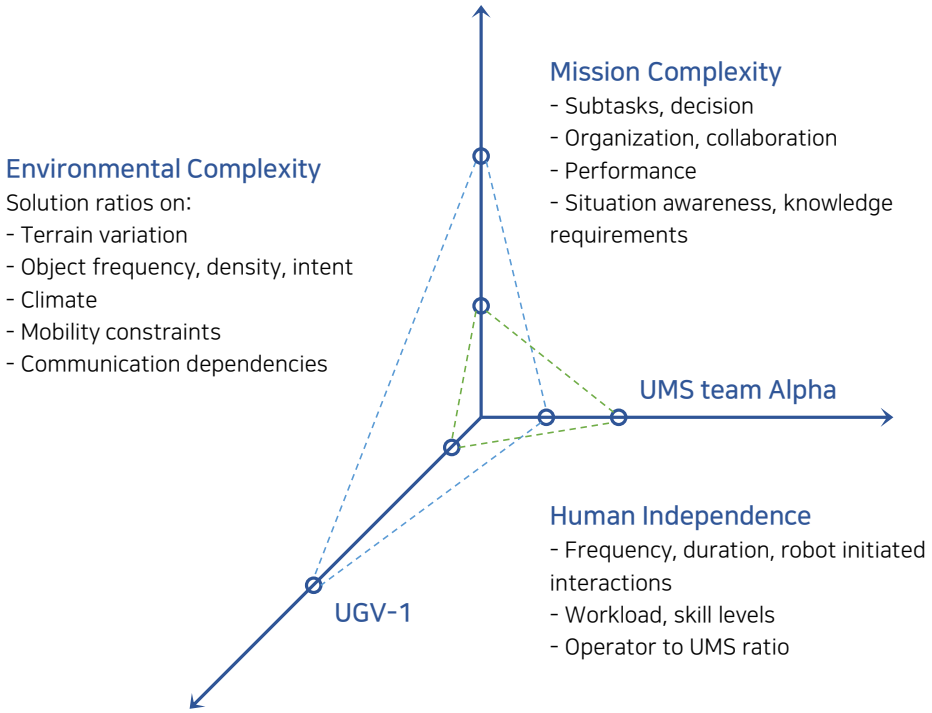

Figure 1. Three-dimensional structure of the ALFUS framework: mission complexity, environmental complexity, and human independence (adapted from Huang et al., 2007 [7]). Note: Dashed lines connect each example system's assessed levels on the three axes, forming the autonomy profiles of UGV-1 (blue dashed) and UMS Team Alpha (green dashed).

UGV-1 represents a configuration characterized by low environmental complexity and high operator involvement, whereas UMS Team Alpha represents highly autonomous operations under high mission complexity and demanding environmental conditions with relatively limited operator involvement. The ALFUS framework approach expands autonomy from a simple measure of technological level into a performance-oriented evaluation framework that reflects operational conditions. It later influenced the development of Draper's mobility–acquisition–protection (MAP) model and the US Air Force Research Laboratory's (AFRL's) ACL model.

The MAP model, originally proposed by Hasslacher and Tilden at Los Alamos National Laboratory, applies the concept of biological survival to autonomy assessment and evaluates autonomy using three functional dimensions: mobility, acquisition, and protection [11]. By examining the independent operation of a system in relation to survival and mission continuity, this model assesses survivability characteristics associated with autonomy.

The ACL model proposed by Bruce Clough at AFRL extends the OODA loop into a multidimensional autonomy framework and classifies autonomy into eleven levels (Levels 0–10) using three dimensions: perception and situational awareness, analysis and decision-making, and communication and cooperation [25]. By evaluating autonomous capability across these OODA-based dimensions and integrating the results into an overall autonomy measure, ACL provides a function-oriented but operationally grounded assessment of unmanned systems. Figure 2 presents a comparison of Draper's MAP model (left) and AFRL's ACL model (right), showing how autonomy assessment evolved from a function-oriented perspective toward an operations-oriented perspective.

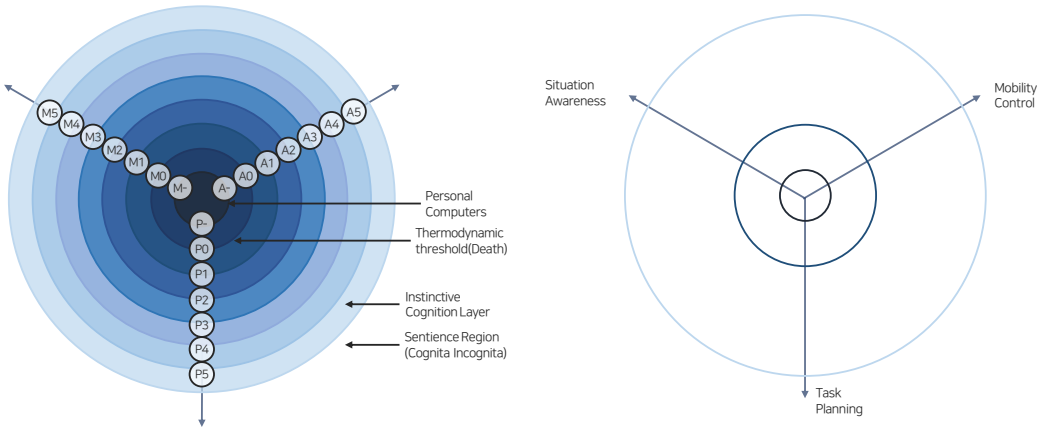

Figure 2. Conceptual comparison of the MAP model (**left**) and the ACL model (**right**) (adapted from Hasslacher and Tilden, 1995 [11]; Clough, 2002 [25]).

Li et al. proposed the Pyramid Chart for autonomy evaluation [14]. This model evaluates autonomy using four dimensions: human–machine interface (HMI), situational awareness, environmental adaptability, and decision-making capability. The area formed by the four dimensions represents the overall system autonomy level and enables intuitive comparison across systems. Wang and Liu later extended this idea into the Cobweb model, wherein multiple evaluation dimensions are arranged radially to form a multidimensional autonomy profile [15]. This visualization approach supports flexible expansion of evaluation factors and comparative analysis across systems. Figure 3 presents a comparison of Li et al.'s Pyramid Chart (left) with Wang and Liu's Cobweb model (right).

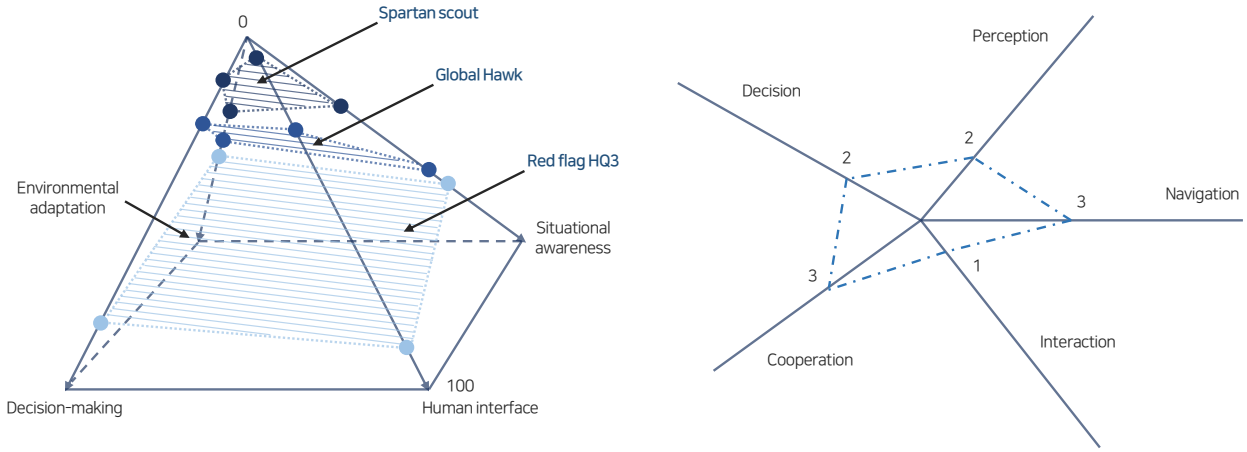

Figure 3. Autonomy evaluation concepts: the Pyramid Chart (**left**) and the Cobweb model (**right**) (adapted from Li et al., 2012 [14]; Wang and Liu, 2012 [15]).

The aforementioned models advanced the field by representing autonomy as a multidimensional structure and shifting LOA research from qualitative classification toward quantitative comparison. However, two limitations persist from a human–machine interaction perspective. First, although these schemes are multidimensional, they typically collapse the dimensions—including human independence and operator involvement—into a single aggregate autonomy index, so that the interaction is not resolved by mission function and a weakly autonomous, interaction-heavy function can be masked by high scores elsewhere. Second, they still treat autonomy mainly as a measure of technical maturity or functional capability and do not explain how changes in the human–machine division of labor propagate through MOPs to MOEs. Consequently, the role of LOA as an operational factor that shapes mission performance through human–machine interaction remains insufficiently explained even in multicriteria approaches. Table 4 consolidates this comparative analysis by grouping the reviewed schemes according to their shared limitations and identifying the prospective directions for overcoming each.

Table 4. LOA Classification Schemes Grouped by Shared Limitation, with Prospective Directions.

Scheme Group (Representatives)	Shared Limitation	Prospective Direction
Operator-involvement scales: Sheridan [8]; Parasuraman et al. [9]; SAE J3016 [10]; Moray et al. [17]; AIAA and AAD hands-on-time frameworks [4,16]	Autonomy interpreted as reduced human intervention along a single system-wide scale; where generic interaction stages are distinguished, autonomy is still not decomposed by mission function	Represent autonomy as a function-specific condition rather than a single platform attribute
Mission-oriented single-axis extensions: UCAV and FCS programs [5]; NIAG SG-75 [12]; Suresh and Ghose [13]	Mission and communication criteria added, but autonomy still classified along one technical progression; operational pathways from LOA to MOPs and MOEs unspecified	Link LOA explicitly to the MOP–MOE hierarchy through an LOA–MOP–MOE structure
Multicriteria frameworks: ALFUS [7]; MAP [11]; ACL [25]; Pyramid Chart [14]; Cobweb model [15]	Multidimensional autonomy profiles are not retained as mission-function-specific performance-shaping variables in effectiveness analysis; the HMT dimension and the LOA–MOP–MOE propagation pathway remain under-specified	Preserve function-level autonomy axes and treat human–machine teaming as an explicit, performance-shaping function

This limitation provides the motivation for the framework proposed in the following sections.

3. ME Framework and Prior Studies on Weapon System Effectiveness

3.1. DoD ME Framework for Effectiveness Evaluation

The ME Guide (2020) presents a standardized framework for analyzing complex systems and systems of systems from a mission perspective and linking the results to acquisition and investment decisions [19]. In this framework, ME decomposes missions into functional elements, integrates the operational and system capabilities required for mission execution, and uses data-driven models and simulations to evaluate alternative system configurations. Moreover, effectiveness evaluation follows a hierarchical structure. COIs are first derived from mission objectives and operational concepts. MOEs are then defined to evaluate mission accomplishment, and MOPs are specified as system-level indicators that support each MOE. The relationships among COI, MOE, and MOP are maintained

consistently across mission scenarios, enabling traceable and comparable effectiveness analyses across alternative system configurations.

Table 5 presents a summary of the common MOE items used in the ME framework. In practice, weapon system evaluations do not apply all MOE items uniformly. Instead, evaluators select indicators according to mission characteristics and the evaluation phase, including requirements definition, test planning, alternative analysis, and acquisition decision-making. Moreover, the ME Guide does not define autonomy as an independent evaluation dimension or explicitly connect it to specific classification schemes, such as LOA or ACL. In most cases, autonomy is indirectly reflected through scenario assumptions and selected MOPs. Given that the structure and operational role of autonomy differ across studies, a review of prior research on its incorporation into MOE–MOP frameworks is presented in the following section. Notably, several of these categories—in particular readiness and cost—express a reliability, availability, maintainability, and cost (RAM-C) perspective on effectiveness, alongside the mission-satisfaction perspective. Effectiveness evaluation is therefore most complete when it considers both a mission perspective and a RAM-C (supportability) perspective, a duality that is equally recognized in weapon-system ME and in the safety and reliability engineering of automated driving [26].

Table 5. Common MOE Categories in the ME Framework (Mission Engineering Guide, 2020 [19]).

Category	Key Contents
Mission satisfaction	Achievement of the final mission objective
Losses	Loss of equipment or personnel due to enemy attacks
Expenditures	Amount of expendable assets used during the operation (e.g., number of munitions employed for message delivery or target destruction)
Cost	Cost of developing alternative systems, technologies, or concepts and cost of conducting operations
Time	Total time required to complete the mission
Repetition	Number of repetitions required to meet the mission satisfaction criteria
Readiness	Level of force readiness for immediate combat deployment
Uncertainty	Level of uncertainty associated with the aforementioned items
Return on investment	Ratio of one item to another (e.g., ratio of mission satisfaction to expenditures)

3.2. Comparative Review of MOE–MOP Structures and Autonomy in Prior Weapon System Studies

Based on the broad set of MOEs suggested in the ME guide, most existing studies adopt three indicators, namely, mission satisfaction, losses, and return on investment (ROI), as primary outcome measures. To improve comparability and decision relevance, the review in this section focuses on these three indicators; the mission-level MOE instantiated in the present case study combines the time and mission-satisfaction categories of Table 5. Prior studies were reviewed to identify system- and function-level MOPs involved in each scenario, including detection and communication, fire or neutralization, and maneuver. The relationships between MOPs and MOEs were then compared across studies. Notably, many studies evaluated effectiveness assuming that unmanned or MUM-T systems operate at the highest autonomy level. Table 6 presents a summary of the MOE–MOP structures and evaluation scenarios used in prior studies.

Table 6. MOE and MOP Components Used in Prior Studies by Weapon System Category.

Category	Mission Effectiveness	Losses	ROI
KAAV + UAV + UGV (Kim and Lee, 2022 [20])	Enemy survival rate: $ESR = \frac{F_E}{F_0} \times 100\,(\%)$ (E_E: surviving force; E_0: initial friendly and enemy forces)	Friendly force survival rate: $FSR = \frac{F_E}{F_0} \times 100\,(\%)$ (F_E: surviving force; F_0: initial enemy forces)	Evaluating changes in combat effectiveness as the number of UGVs and UAVs increases
USV (Park et al., 2022 [21])	Presenting the number of enemy vessels that can be hit based on enemy naval gun hit rate, the number of approaching friendly USVs, and the number of guided rockets carried by each friendly USV	Presenting the survival rate of friendly USVs according to changes in naval gun hit rates of North Korean patrol boats, Japanese destroyers, and Chinese destroyers	Presenting the survival rate of friendly USVs as the number of friendly USVs and friendly hit rate increase
Army TIGER (Yoon et al., 2024 [22])	Mission success rate for each capability category—identification, survivability, penetration, and strike: $MS_{AT} = \frac{\text{Mission success force}}{\text{Total force}} \times 100\,(\%)$	$S_{AT} = \frac{\text{Surviving force}}{\text{Total force}} \times 100\,(\%)$	–
Category	**Mission Effectiveness**		
UGV (Lee et al., 2014 [23])	Mine detection: $MOE(mine) = \frac{P_d(mine)}{T}$ (P_d: number of detected mines; T: detection time). CBRN detection: $MOE(nbc) = \frac{P_d(nbc)}{T+t'}$ (nbc: CBRN; T: detection time; t': contamination type). Search and reconnaissance: $MOE(Rec) = \sum_{i=1}^{n} \left[\frac{P_i(plan)+P_i(RS)}{T_i+t_i}\right]$ ($P_i(plan)$, $P_i(RS)$: detection probability on planned route and in RS area; T_i, t_i: detection time on planned route, in RS area). Casualty evacuation: $MOE(rescue) = \frac{M}{T}$ (M: number of casualties rescued; T: total activity time spent on casualty evacuation). Firing mission: $MOE(fire) = \sum_{i=1}^{n} \left[P_i(hit)\,\omega_i E_i\right]$ ($P_i(hit)$: hit probability; ω_i: probability value of the damage level; E_i: effect value of the damage level).		
Interception-centered integrated system (Shin et al., 2019 [24])	Sensor system detection rate: $MS_{AT} = \frac{\text{Mission success force}}{\text{Total force}} \times 100\,(\%)$. Weapon system hit rate: $P_h = 1 - \exp\left(-0.6931 \times \frac{R^2}{CEP^2}\right)$ (R: target size; CEP: circular error probable). Platform destruction rate: $P_k = 1 - \theta^{\left(P_d P_h P_{sys} R_w / T_{strength}\right)^2}$ (P_d: sensor detection rate; P_h: weapon system hit rate; P_{sys}: combat system reliability; R_w: weapon system reliability; $T_{strength}$: target hardness).		

Kim and Lee (2022) [20] used a scenario-based simulation to compare the operational effects of the Korean Assault Amphibious Vehicle (KAAV), UAVs, and UGVs in amphibious operations. They evaluated effectiveness by linking mission-level indicators with system-level performance indicators. Mission satisfaction and losses were used as the primary MOEs, with mission satisfaction defined as the inverse of the enemy survival rate and the losses represented by the friendly survival rate. The MOPs were organized according to platform function: firepower and maneuver for the KAAV and detection and communication for UAVs and UGVs. Under this structure, surveillance, maneuver, and firepower were linked sequentially to mission-level outcomes. The simulation results showed that prior reconnaissance and information sharing reduce breaching time and casualties and that bottlenecks in surveillance, maneuver, and firepower strongly influence mission performance. However, the analysis assumed a high LOA and did not explicitly represent autonomy-related factors, such as operator involvement, decision latency, replanning, or cooperative efficiency within the MOP structure.

Park et al. (2022) [21] compared combat effectiveness across three engagement scenarios wherein a friendly unmanned surface vessel (USV) confronted a North Korean patrol craft, a Japanese destroyer, and a Chinese destroyer, respectively. They used mission satisfaction, losses, and ROI as the main MOEs. Mission satisfaction was represented by engagement performance, losses by USV survival rate, and ROI by the relationship between survivability, force size, and hit probability. The main MOPs comprised firepower accuracy variables, such as naval-gun hit probability and guided-rocket hit probability. Improvements in firepower accuracy increased mission performance and survivability, and the MOEs were highly sensitive to opponent type, force size, and hit probability. However, autonomy was not explicitly modeled in the analysis. Consequently, the study effectively assumed highly autonomous operations and did not represent autonomy-sensitive MOPs, such as C2 resilience, distributed maneuver, or operator involvement. Thus, identifying how changes in LOA propagate to MOEs through firepower-centered MOP structures remains difficult.

Yoon et al. (2024) [22] used AWAM to evaluate the combat effectiveness of an Army TIGER battalion-level force. They applied a two-factor design that compared effectiveness before and after (a) Army TIGER integration and (b) drone deployment. The MOEs

were defined as mission success rates by capability category (identification, survival, advance, and fire) together with the friendly force survival ratio. The MOPs comprised basic performance variables, such as maneuver speed, identification rate, line-of-sight range, and concealment ratio. This capability-based structure helps identify the functions that contribute most strongly to mission outcomes. However, because the relationships between individual MOPs and MOEs were not explicitly modeled, determining how changes in performance variables influence mission success rates remains challenging.

Lee et al. (2014) [23] analyzed the operational effectiveness of UGVs by classifying missions into mine detection, chemical, biological, radiological, and nuclear detection, search and reconnaissance, rescue, and fire missions. Moreover, they proposed separate MOE formulations for each mission type using relevant performance indicators. Furthermore, they used target detection probability as a common performance indicator and extended it to target hit probability in fire missions. This mission-specific formulation is useful because it shows that different mission types require different effectiveness measures. However, the study mainly focused on mission-specific MOE formulation and provided limited discussion of how operational factors interact with performance measures across different mission conditions.

Shin et al. (2019) [24] analyzed guided-weapon-based coastal defense by conducting 100 simulation runs across different combinations of guided rockets. They defined (a) attack and landing success rates as the main MOEs and (b) engagement-related measures, such as detection rate, hit rate, destruction rate, combat system processing time, and firing interval, as the MOPs. Differences in the guidance method and weapon range considerably influenced mission effectiveness through fire-related MOPs. However, the analysis centered only on weapon performance variables and gave limited attention to how operational factors interact with effectiveness measures.

As demonstrated by the preceding review, most existing effectiveness evaluations do not explicitly consider autonomy and often assume a fixed or uniformly high LOA. Two gaps follow. First, autonomy is treated as a single system-wide attribute rather than a function-specific condition that matures unevenly across surveillance, maneuver, fire or neutralization, and command and control. Second, and more fundamentally, the human–machine interaction that autonomy reshapes—operator involvement, workload, supervisory control, and the sharing of control between the human and the system—is rarely represented within the MOP–MOE structure, even though several of the reviewed studies note operator involvement as an unmodeled factor. Consequently, the interaction bottlenecks that arise when control is partially transferred to unmanned vehicles remain invisible to these evaluations. Addressing both gaps requires MOP and MOE structures that represent autonomy as a function-specific vector and treat human–machine teaming as an explicit, performance-shaping function rather than a background assumption.

4. A Function-Specific Human–Machine Interaction Framework for Effectiveness Analysis

This section proposes a function-specific human–machine interaction (HMI) framework for effectiveness analysis that addresses the two gaps identified in the preceding review: (i) the treatment of LOA as a single, system-wide scalar, and (ii) the omission of human–machine interaction—operator workload, supervisory control, and the human–machine division of labor—from the MOP–MOE structure. The framework retains the COI–MOE–MOP structure of mission engineering but defines LOA by function and treats it as a time-varying condition, so that autonomy shapes mission effectiveness through an explicit LOA–MOP–MOE pathway rather than through a single system-level assumption.

The framework operationalizes the function-specific view of human–machine interaction introduced in Section 2. In a MUM-T mission, the domain-neutral stages of perception, assessment, decision, and action are instantiated as concrete mission functions: surveillance, maneuver, fire or neutralization, and inter-asset command and control (C2). To these we add human–machine teaming (HMT) as an explicit, first-class function that captures the operator–system division of labor itself, including hands-on time, workload, and supervisory structure. Human–machine interaction in this framework is therefore not a single channel but a property distributed across all five functions, each of which carries a different facet of it: LOA_S governs how perception and situational awareness are shared; LOA_{C2} governs supervisory coordination across multiple assets; LOA_F governs the operator's direct control and approval of engagement or neutralization; LOA_M governs how often the operator must intervene in movement; and LOA_{HMT} measures the operator–system division of labor most directly, through hands-on time, workload, and supervisory structure. Because the interaction is borne jointly by these axes rather than by a single index, the framework can expose interaction bottlenecks—functions whose low autonomy or heavy operator dependence constrains overall effectiveness even when other functions are highly autonomous—and locate them at the level of the responsible function.

4.1. Definition of Function-Specific LOA and the Time Index

Existing effectiveness evaluations tend to either treat LOA as a single scalar value applied to entire weapon systems or, for analytical convenience, assume that systems operate at the highest autonomy level throughout. However, in MUM-T systems, autonomy matures differently across operational functions, such as surveillance, maneuver, fire or neutralization, C2, and HMT, and mission outcomes are often determined by bottlenecks within these interconnected functions. Accordingly, the present framework defines autonomy not as a single system-level value but as a function-specific autonomy vector within the COI–MOE–MOP structure of ME:

$$\mathbf{LOA}(t) = \big(LOA_S(t),\, LOA_M(t),\, LOA_F(t),\, LOA_{C2}(t),\, LOA_{HMT}(t)\big). \tag{1}$$

- LOA_S—surveillance, detection, and identification autonomy (including situational awareness).
- LOA_M—maneuver and navigation autonomy (path planning, avoidance, and replanning).
- LOA_F—fire or neutralization autonomy (target selection, fire control, engagement, and neutralization management).
- LOA_{C2}—inter-asset C2 autonomy (link independence, resilience, and distributed cooperation among assets).
- LOA_{HMT}—human–machine autonomy (hands-on time, workload, and supervisory structure between operator and system).

Because LOA_{HMT} is the axis that most directly represents the operator–system division of labor, it is defined operationally rather than as an abstract level. Following the human–autonomy teaming measurement literature [27], LOA_{HMT} is assessed along three observable proxies: (i) the operator's hands-on control time as a fraction of mission time, consistent with hands-on-time definitions of autonomy [16]; (ii) operator workload, measured or predicted with a standard instrument such as the NASA Task Load Index [28]; and (iii) the achievable span of control, i.e., the number of unmanned assets one operator can effectively supervise [29]. A configuration dominated by continuous teleoperation—high hands-on time, high workload, and a span of control of at most one asset—corresponds to a low LOA_{HMT} and a reduced human–machine teaming contribution, whereas a configuration in which the operator supervises and approves autonomous action—low hands-on time, moderate workload, and a span of control greater than one—corresponds to a higher

LOA_{HMT}. This operational definition makes LOA_{HMT} estimable from mission data or structured expert judgment rather than assigned by assumption. The dominant control mode provides the primary classification, and hands-on time, workload, and span of control serve as corroborating indicators; when the indicators disagree, a prespecified conservative rule or a structured expert-judgment procedure is applied. The framework does not fix the number of ordinal levels for LOA_{HMT}; the granularity is chosen per application, and the mapping used for the staged alternatives of the present case is given in Section 5.3.

This five-function decomposition is the instantiation appropriate to the MCM mission rather than a fixed taxonomy. The framework requires only that autonomy be expressed as a vector of function-specific, interaction-bearing axes that propagate to the MOE; the particular axes are tailored to the mission and platform under study. Functions may be added, merged, or omitted accordingly—for example, a maritime surveillance task may drop LOA_F and refine LOA_S, whereas a strike task may elaborate it—while the human–machine teaming axis LOA_{HMT} remains across instantiations because some division of labor between operator and system is always present. This flexibility is what allows the same structure to be carried beyond the present case, a point taken up in Section 6.

Furthermore, t does not represent calendar time but a discrete index representing stages of autonomy evolution. Autonomy does not increase continuously over time; rather, it changes stepwise through events, such as sensor upgrades, algorithm updates, software improvements, operational data accumulation, and improvements in the HMI. Accordingly, autonomy at time step $t+1$ is defined as

$$LOA_i(t+1) = LOA_i(t) + \Delta LOA_i(t), \tag{2}$$

where $\Delta LOA_i(t)$ represents the autonomy level change of function i at time step t due to technological upgrades or operational learning effects.

4.2. Function-Specific LOA–MOP–MOE Linkage and Conditional MOPs

As discussed earlier, LOA is defined as a function-specific vector whose maturity varies across operational functions. In the proposed framework, LOA serves as an operational condition that shapes the MOPs associated with each function. Accordingly, the function-specific MOP at time step t can be expressed as

$$MOP_j(t) = f_j\big(\mathbf{LOA}(t),\, \mathbf{Env}(t),\, \mathbf{X}(t)\big). \tag{3}$$

- $\mathbf{LOA}(t)$—function-specific autonomy vector.
- $\mathbf{Env}(t)$—environmental conditions, including mission difficulty, battlefield conditions, and threat level.
- $\mathbf{X}(t)$—intrinsic technical performance inputs of system components (e.g., reference platform speed, probabilistic performance, processing capacity, and component efficiency coefficients).

In the aforementioned formulation, $\mathbf{X}(t)$ represents the intrinsic technical capabilities of the system components, which may evolve together with autonomy through improvements in sensing, processing, communication, and cooperative control technologies. Accordingly, $\mathbf{LOA}(t)$ and $\mathbf{X}(t)$ should not be interpreted as fully independent variables. Instead, the framework distinguishes them analytically to isolate the operational effects of autonomy from the underlying technical capabilities of the system. Similarly, $\mathbf{Env}(t)$ conditions both the realized intrinsic performance and the effective autonomy of each function—as reflected in the environmental-complexity dimension of autonomy schemes such as ALFUS [7]—so the three inputs are mutually interdependent; the framework separates them analytically to isolate the operational effect of autonomy. In the present case,

$\mathbf{Env}(t)$ is treated as an exogenous input and held constant across alternatives to isolate the effects of the function-specific LOA configuration. Depending on function, autonomy effects may be incorporated through (i) LOA-dependent multipliers on intrinsic performance, (ii) intrinsic values such as detection probabilities, or (iii) both. Therefore, MOP should be interpreted not as a fixed performance value but as a conditional performance jointly shaped by autonomy, operational environment, and technical capability.

The operational implications of this conditional MOP structure can be illustrated as follows. Consider two configurations that report a detection probability of 0.7. Under a low-LOA condition with human-centric search, heavy operator involvement and limited persistent surveillance may increase missed coverage and the likelihood of repeated search operations. Under a high-LOA condition with highly autonomous search, automatic path planning, persistent surveillance, and re-search capability can maintain the same nominal detection probability more consistently while enabling multiplatform cooperation and wider search coverage.

Thus, even when the same performance indicator is reported numerically, its operational meaning may differ depending on the LOA configuration. These differences influence how functions are sustained, coordinated, and repeated during mission execution, ultimately producing different MOE outcomes (e.g., mission success rate, completion time, losses, and ROI). In the proposed framework, these conditional MOPs may appear at different analytical levels, including function-level and mission-process MOPs, depending on the mission execution stage. Reflecting this structure, MOE can be represented as

$$MOE(t) = g\big(MOP_1(t), \ldots, MOP_n(t) \mid \mathbf{LOA}(t)\big). \tag{4}$$

Under this formulation, MOE is not interpreted as a simple aggregation of MOP values but as an outcome conditioned by the LOA structure under which these MOPs are realized. This differs from prior studies that implicitly fixed LOA or assumed the highest autonomy level throughout evaluation. By defining LOA as a function-specific vector, the proposed framework recognizes that operational functions mature at different rates and that deficiencies in individual functions may become mission bottlenecks. Likewise, by interpreting MOPs as conditional performances jointly shaped by LOA, operational environment, and intrinsic system capability, the framework explains how identical numerical performance values may yield different operational outcomes depending on persistence, cooperability, and scalability during mission execution.

Operationally, the framework is structured into five analytical layers: (i) intrinsic system performance inputs, (ii) LOA-dependent performance coefficients, (iii) conditional functional MOPs, (iv) mission-process MOPs, and (v) mission-level MOEs. Function-specific LOA first modifies the performance coefficients, such as cooperation, interface, and maneuver efficiency. These coefficients then shape the functional MOPs, including detection probability, effective search speed, and effective cooperative mission-unit capacity. The functional MOPs aggregate into mission-process MOPs, such as search time and neutralization time, and the mission-process MOPs ultimately propagate into mission-level MOEs. Figure 4 summarizes this five-layer structure, the three determinants of conditional performance, and the origin of the function-specific interaction bottleneck. Section 5 presents the application of this five-layer structure to a minehunter-centered MCM case.

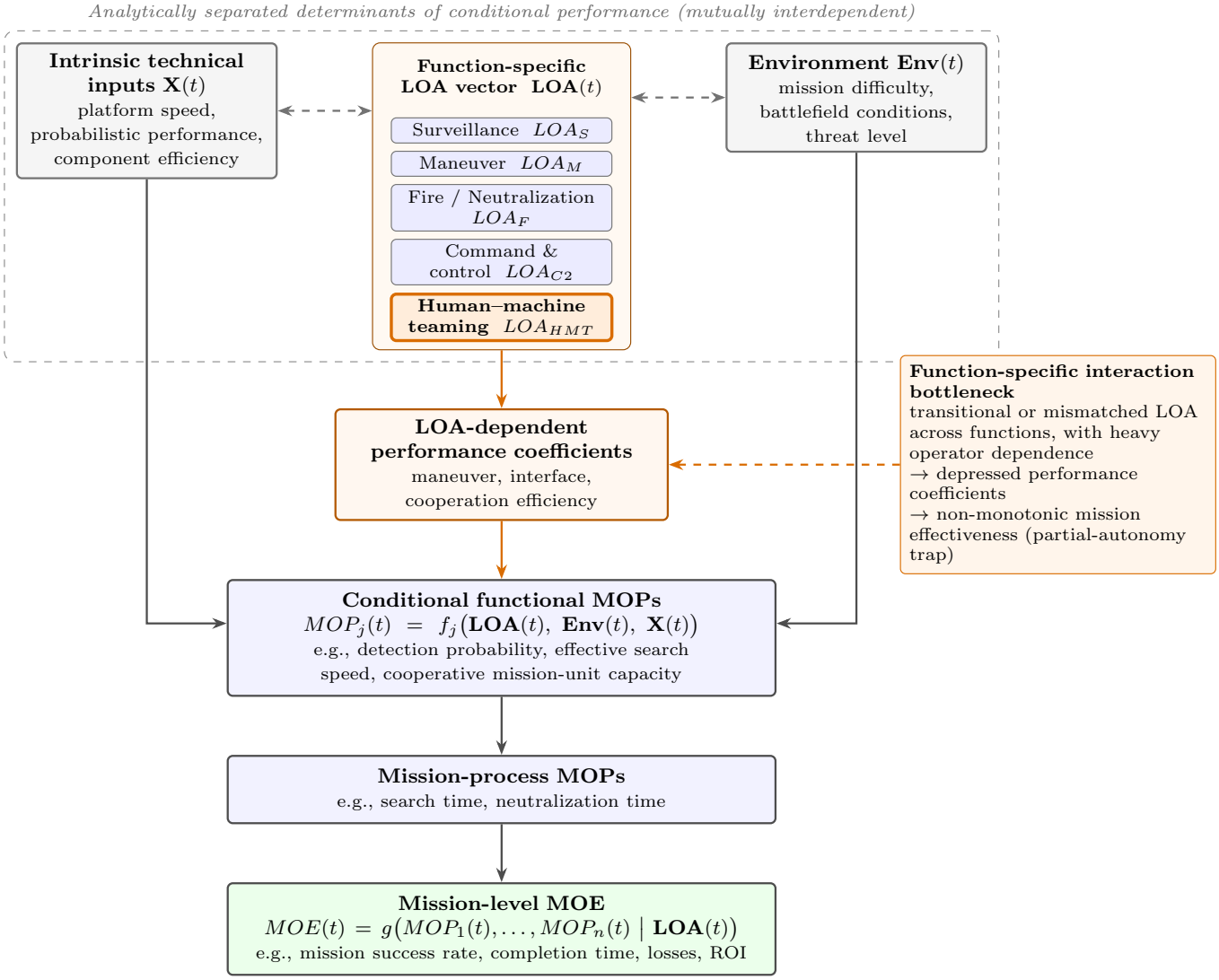

Figure 4. Five-layer structure of the proposed function-specific effectiveness framework. The three determinants of conditional performance—the intrinsic technical inputs $\mathbf{X}(t)$, the function-specific LOA vector $\mathbf{LOA}(t)$, and the operational environment $\mathbf{Env}(t)$—are mutually interdependent but analytically separated to isolate the operational effect of autonomy. Function-specific LOA acts through the LOA-dependent performance coefficients, which jointly with $\mathbf{X}(t)$ and $\mathbf{Env}(t)$ form the conditional functional MOPs and propagate through mission-process MOPs to the mission-level MOE. The callout marks the origin of the function-specific interaction bottleneck underlying the partial-autonomy trap.

5. Application to a Minehunter-Centered MCM Case

5.1. Case Setup

Based on the development direction of the Republic of Korea Navy's next-generation minehunter (MSH-II) and mine-warfare combat system, this case considers an operational concept wherein a manned minehunter serves as a mothership and C2 platform. Under this concept, mine search equipment, mine disposal vehicles (MDVs), underwater and surface unmanned assets, and the mine-warfare combat system are integrated stepwise [30].

The operational scenario involves the rapid opening of a priority transit channel within a port-approach lane. The objective is not exhaustive mine clearance across the entire threat area but the timely opening of an operationally usable channel under binding time constraints. The search area is defined as a single channel of 5 nm × 0.5 nm, and the

operational flow comprises two stages: search and individual neutralization. In this scenario, the case examines staged configurations A0–A3, which represent successive developmental stages of a minehunter-centered MCM system toward MUM-T integration.

Each alternative is specified as a representative configuration corresponding to a distinct developmental stage in autonomy and system integration. For each alternative, $\mathbf{LOA}_j \equiv \mathbf{LOA}(t_j)$, $\mathbf{Env}_j \equiv \mathbf{Env}(t_j)$, and $\mathbf{X}_j \equiv \mathbf{X}(t_j)$ are defined and the case analysis uses the alternative index $j \in \{A0, A1, A2, A3\}$ instead of the generic time index t. The alternatives capture stepwise changes not only in LOA but also in mine search capability, neutralization capability, and combat system integration.

This case aims not to forecast the performance of a specific MCM system, but to demonstrate how the proposed function-specific LOA–MOP–MOE structure can be instantiated under staged MUM-T configurations. The operational concept and asset configuration of each alternative are summarized in Table 7.

Table 7. Analytical Alternatives (A0–A3): Operational Concepts and Asset Configuration.

Alt.	Operational Concept	Mothership	Search Configuration	Neutralization Configuration
A0	Conventional manned-centric minehunter	Two legacy MHC/MSH ships	Hull-mounted sonar, manned	Explosive Ordnance Disposal (EOD) divers, manned
A1	Current minehunter with remote-controlled MDVs	Two current MSH ships	Hull-mounted sonar, manned	Remote-controlled MDVs
A2	Initial MSH-II configuration	Two MSH-II ships	Hull-mounted sonar and autonomous search vehicles	Semiautonomous MDV-IIs
A3	Advanced MSH-II concept	One MSH-II ship	Networked unmanned underwater vehicle (UUV)/USV search assets	Operator-approved autonomous neutralization systems

Note: For the case calculation, N_j^{search} is set to 2 for A0 and A1 and 4 for A2 and A3, whereas $N_j^{neutralize}$ is set to 2 for all alternatives. The increase from 2 to 4 from A2 onward reflects the introduction of autonomous search vehicles that operate in parallel with the mothership's hull-mounted sonar.

5.2. Model Formulation for the MCM Case

The MCM case instantiates the five-layer structure introduced at the end of Section 4. In this case, the mission-level MOE, i.e., timely route-opening capability, is linked to mission-process MOPs, such as search time and neutralization time. These MOPs are derived from functional MOPs, including effective search speed, effective neutralization capacity, and stage-specific effective cooperative mission-unit capacity. The functional MOPs are jointly formed by LOA-dependent performance coefficients (η_M, η_F, η_{C2}, η_{HMT}^{stage}), system-specific intrinsic technical inputs X_j, and mission-scenario inputs Env_j. The variables and functions used in the case are mapped to the five-layer structure in Table 8. In this MCM case, the general fire or neutralization dimension LOA_F is instantiated as mine neutralization autonomy.

The mission-level MOE for the present case, i.e., MOE_{Open}, measures the timely route-opening capability under required quality conditions:

$$MOE_{Open,j} = I_{Q,j} \cdot \min\left(1, \frac{T_{req}}{T_{MCM,j}}\right), \tag{5}$$

where T_{req} denotes the required route-opening time. $I_{Q,j}$ is a quality gate that indicates whether the probabilistic mission requirements are satisfied:

$$I_{Q,j} = \begin{cases} 1 & P_{D,j} \geq P_D^{req},\ P_{I,j} \geq P_I^{req},\ P_{N,j} \geq P_N^{req} \\ 0 & \text{otherwise} \end{cases}, \tag{6}$$

where $P_{D,j}$, $P_{I,j}$, and $P_{N,j}$ are the detection, identification, and neutralization probabilities for the alternative j, respectively, and P_D^{req}, P_I^{req}, and P_N^{req} are the corresponding minimum required thresholds. Equation (5) first applies the quality gate through $I_{Q,j}$ and then assigns a timeliness score using $T_{req}/T_{MCM,j}$, capped at 1. Thus, a quality failure yields zero effectiveness, while a time overrun is penalized proportionally when the quality requirements are satisfied.

The total MCM mission time is decomposed into search and neutralization times:

$$T_{MCM,j} = T_{search,j} + T_{neutralize,j}. \tag{7}$$

This decomposition represents MCM operations as a sequence of functional stages. The search and neutralization times are defined as

$$T_{search,j} = \frac{1}{P_{D,j}} \cdot \frac{L_s R_s}{V_{eff,j}\, N_{eff,j}^{search}} + T_{turn,s}\, N_{turn,s}, \tag{8}$$

$$T_{neutralize,j} = \frac{1}{P_{I,j} P_{N,j}} \cdot \frac{Q}{D_{eff,j}\, N_{eff,j}^{neutralize}}, \tag{9}$$

where L_s is the channel length, R_s is the number of round trips, and Q is the number of objects to be neutralized. The factor $1/P_{D,j}$ adjusts the effective search coverage time to reflect additional search effort caused by detection failure, while $T_{turn,s}\, N_{turn,s}$ is treated as a fixed maneuver overhead. The factor $1/(P_{I,j}P_{N,j})$ adjusts the neutralization time to reflect reconfirmation or reprocessing caused by identification or neutralization failure.

The functional MOPs used in Equations (8) and (9) are $V_{eff,j}$, $D_{eff,j}$, and $N_{eff,j}^{stage}$, which represent effective search speed, effective neutralization capacity, and stage-specific effective cooperative mission-unit capacity, respectively. These MOPs are formed jointly by intrinsic technical inputs, LOA-dependent performance coefficients, and the nominal number of mission units assigned to each stage. For the case calculation, N_j^{search} is set to 2 for A0 and A1 and 4 for A2 and A3, whereas $N_j^{neutralize}$ is set to 2 for all alternatives. Because operator workload and supervisory burden may differ between search and neutralization, the HMT coefficient is defined at the stage level and denoted by $\eta_{HMT,j}^{stage}$. In particular, $\eta_{HMT,j}^{search} = 1.00$ for A0 and A1 because their search configuration remains manned, whereas $\eta_{HMT,j}^{search} = \eta_{HMT}(LOA_{HMT,j})$ for A2 and A3. For the neutralization stage, $\eta_{HMT,j}^{neutralize} = \eta_{HMT}(LOA_{HMT,j})$ for all alternatives. The functional MOPs are defined as follows:

$$V_{eff,j} = V_0 \cdot X_{V,j} \cdot \eta_M(LOA_{M,j}), \tag{10}$$

$$D_{eff,j} = \frac{X_{N,j}}{\tau_0} \cdot \eta_F(LOA_{F,j}), \tag{11}$$

$$N_{eff,j}^{stage} = 1 + (N_j^{stage} - 1) \cdot \eta_{C2}(LOA_{C2,j}) \cdot \eta_{HMT,j}^{stage}. \tag{12}$$

The form of Equation (12) reflects the assumption that the first mission unit is always fully effective, with additional units contributing only to the extent permitted by η_{C2} and η_{HMT}. As η_{C2} and η_{HMT} degrade, the effective number of cooperating units approaches one, even if multiple physical assets are deployed.

Table 8. Five-Layer Structure of the MCM Case.

Layer	Notation	Definition
(1) Intrinsic technical inputs X_j	V_0, τ_0 $X_{V,j}$, $X_{N,j}$ $P_{D,j}$, $P_{I,j}$, $P_{N,j}$	Reference search speed and per-object reference processing time (common reference values); mothership and neutralization-equipment efficiency coefficients (platform-generation effects); detection, identification, and neutralization probabilities of the sensor and neutralization systems.
(2) LOA-dependent performance coefficients	$\eta_M(LOA_{M,j})$, $\eta_F(LOA_{F,j})$, $\eta_{C2}(LOA_{C2,j})$, $\eta_{HMT,j}^{stage}$	Maneuver efficiency, fire or neutralization-stage efficiency, C2 cooperation efficiency, and stage-specific HMT efficiency.
(3) Functional MOPs	$V_{eff,j}$, $D_{eff,j}$, $N_{eff,j}^{stage}$	Effective search speed, effective neutralization capacity, and stage-specific effective cooperative mission-unit capacity.
(4) Mission-process MOPs	$T_{search,j}$, $T_{neutralize,j}$, $T_{MCM,j}$	Search time, neutralization time, and total MCM mission time.
(5) Mission-level MOE	$MOE_{Open,j}$	Timely route-opening capability under required probabilistic mission conditions.

Note: $stage \in \{search, neutralize\}$. Mission-scenario inputs such as L_s, R_s, Q, $T_{turn,s}$, $N_{turn,s}$, and N_j^{stage} are treated as exogenous case inputs and enter Equations (8), (9) and (12).

5.3. Input Values and LOA Settings

The intrinsic technical inputs X_j of each alternative are summarized in Table 9. The reference values $V_0 = 5$ and $\tau_0 = 1$ are common to all alternatives. The platform-generation efficiency coefficients $X_{V,j}$ and $X_{N,j}$, together with the probabilistic performance values $P_{D,j}$, $P_{I,j}$, and $P_{N,j}$, follow notional cumulative step-gain rules:

$$X_{c,j} = (1 + \rho_c)^{k_{c,j}}, \tag{13}$$

$$P_{c,j} = 1 - (1 - P_{c,0})\,\delta_c^{k_{c,j}}, \tag{14}$$

where ρ_c is the per-step efficiency gain, δ_c denotes the per-step residual failure ratio, and $k_{c,j}$ represents the number of generational steps that component c has undergone in alternative j. The present case applies the notional uniform values $\rho = 0.1$ and $\delta = 0.8$. Under this rule, the mothership efficiency coefficient $X_{V,j}$ advances once from A2 onward, whereas the neutralization equipment efficiency coefficient $X_{N,j}$ advances at each stage after A0. The detection and identification probabilities $P_{D,j}$ and $P_{I,j}$ advance from A2 onward, whereas the neutralization probability $P_{N,j}$ advances from A1 onward. The mission-scenario inputs, common to all alternatives, are $L_s = 5\,\mathrm{NM}$, $R_s = 5$, $T_{turn,s}N_{turn,s} = 0.31\,\mathrm{h}$, and $Q = 15$, corresponding to the environmental conditions Env_j in equation (3). These inputs are held constant across alternatives to isolate the effect of function-specific LOA configurations on mission effectiveness.

The LOA-dependent performance coefficients are defined as step functions of the corresponding LOA values, as summarized in Table 10. They follow two main patterns: η_M, representing maneuver efficiency, and η_{C2}, representing C2 cooperation efficiency, increase monotonically with LOA and remain unchanged between LOA 1 and LOA 2. This reflects the assumption that path planning autonomy and inter-asset coordination do not improve substantially during the remote-control stage, but begin to improve when integrated combat system support and cooperative autonomy are introduced from LOA 3 onward.

Table 9. Intrinsic Technical Inputs and Probabilistic Performance Values by Alternative.

Alt.	$X_{V,j}$	$X_{N,j}$	$P_{D,j}$	$P_{I,j}$	$P_{N,j}$
A0	1.000	1.000	0.650	0.600	0.700
A1	1.000	1.100	0.650	0.600	0.760
A2	1.100	1.210	0.720	0.680	0.808
A3	1.100	1.331	0.776	0.744	0.846

Table 10. LOA-Dependent Performance Coefficients (Step Values and Characteristics).

Coefficient	LOA 1	LOA 2	LOA 3	LOA 4	Characteristic
η_M	1.00	1.00	1.07	1.13	Monotonic increase (flat between LOA 1 and 2)
η_F	1.00	0.70	1.20	1.60	Nonmonotonic—valley at LOA 2
η_{C2}	0.65	0.65	0.88	1.00	Monotonic increase (flat between LOA 1 and 2)
η_{HMT}	1.00	0.40	0.70	0.85	Nonmonotonic—valley at LOA 2

Note: η_{HMT} represents the base HMT performance coefficient associated with each LOA value. In Equation (12), the realized HMT coefficient is applied at the mission-stage level and denoted by $\eta^{stage}_{HMT,j}$.

Furthermore, η_F (representing fire or neutralization-stage efficiency) and the base HMT coefficient η_{HMT} exhibit a transient drop at LOA 2 before recovering at LOA 3 and improving further at LOA 4. This pattern is referred to as the LOA-2 valley. This valley reflects the teleoperation penalty introduced during the transition from manual operation to remote-controlled operation, including increased operator workload and limited supervisory scalability. In the case formulation, the realized HMT coefficient is applied at the mission-stage level as $\eta^{stage}_{HMT,j}$ using the step values in Table 10 according to the stage (search or neutralization). This nonmonotonic structure is the primary reason that A1 shows a longer neutralization time than A0 and is central to the interpretation of the case results.

The nonmonotonic structures of η_F and η_{HMT} in Table 10 are motivated by two related but distinct human-factors mechanisms. Under continuous teleoperation, the operator remains directly in the control loop but faces near-continuous control demands, mediated perception, interface burden, and control latency; these demands increase workload, reduce task efficiency, and limit the number of assets that one operator can manage effectively—so much so that a single teleoperated asset may require more than one operator [29,31]. Under partial or supervisory automation, a different mechanism arises: reduced continuous engagement can degrade situation awareness and impair the operator's ability to re-enter the control loop during failures, exceptions, or control transitions—the out-of-the-loop performance decrement [32–34]. Meta-analytic evidence further identifies partial automation as a problematic transition zone in which emergency-response performance is worse than under manual control while workload remains higher than under high automation [35]. In the present case, the LOA-2 valley corresponds to the first mechanism—the continuous-teleoperation stage of A1—whereas the second mechanism explains why supervisory autonomy at higher LOA relieves workload and restores span of control without eliminating the operator's residual monitoring burden [27]; the incomplete recovery of η_{HMT} (0.70–0.85 rather than 1.00) reflects this residual supervisory cost.

The drop of η_F at LOA 2 reflects the same teleoperation penalty at the level of the neutralization task itself. Holding the neutralization equipment fixed, remotely operating a disposal vehicle is less efficient at the task than direct execution, because teleoperation removes the direct haptic and sensory feedback available to a diver and adds control latency and restricted perception. A meta-analysis of teleoperation systems shows that such direct feedback significantly improves task completion time and accuracy, so its absence degrades manipulation performance [31], an effect further compounded by control

delay [29]. As operator-supervised semiautonomy (LOA 3) and operator-approved autonomy (LOA 4) restore automated assistance to the neutralization task, η_F recovers and exceeds the manual baseline. Here η_F denotes the per-task efficiency of the neutralization action under a given control mode with equipment held fixed; it does not represent the safety or force-protection benefits of removing personnel from the minefield, which are real but lie outside the mission-time MOE modeled here. The specific step magnitudes in Table 10 ($\eta_F = 0.70$, $\eta_{HMT} = 0.40$ at LOA 2) are notional values chosen to be consistent with the direction and stage-specificity established by this literature.

In interpreting Table 10, LOA_{HMT} and η_{HMT} serve different analytical purposes. LOA_{HMT} is an ordinal category describing the dominant operator–system role-allocation mode, classified from the observable proxies defined in Section 4.1, whereas η_{HMT} is a separately calibrated continuous multiplier representing the mission-performance contribution realized under that role allocation. The proxies support the classification of LOA_{HMT}; they do not numerically determine η_{HMT} by themselves. Table 11 lists the operator–system role-allocation modes corresponding to the LOA_{HMT} values used in Tables 10 and 12. The granularity follows the four staged configurations of the case, and the modes consolidate role-allocation distinctions that recur across the taxonomies reviewed in Section 2 rather than introducing a new scale.

Table 11. Role-Allocation Modes Corresponding to the LOA_{HMT} Values Used in This Case.

LOA_{HMT}	Dominant Role Allocation	Corresponding Constructs in Reviewed Schemes
1	Direct human execution or manual operation	Manual operation (Sheridan Level 1; UCAV Level 1) [5,8]
2	Continuous teleoperation	Remote control with near-100% hands-on time (AIAA LOA 0; FCS Levels 1–2) [5,16]
3	Supervisory control with human approval of key actions	Management by consent; execution upon operator approval (UCAV Level 2; Sheridan Level 5) [5,8]
4	Goal- and constraint-based collaborative autonomy or management by exception	Management by exception; collaborative operations (UCAV Level 3; AIAA and AAD highest levels) [4,5,16]

The function-specific LOA values for each alternative are summarized in Table 12, reflecting the operational concept of each staged configuration defined in Table 7. The values in parentheses indicate the corresponding base performance coefficients from Table 10, where applicable. LOA_S is included to preserve the function-specific autonomy vector, but search-stage autonomy is reflected by the probabilistic performance values P_D and P_I in Table 9 rather than by a separate η_S coefficient.

A0 represents the conventional manned-centric configuration. With no autonomy systems introduced and EOD divers performing direct neutralization, all functions are set to LOA 1. A1 introduces remote-controlled MDVs at the neutralization stage while retaining the existing minehunter and combat system. In this configuration, LOA_F advances to 2, representing teleoperation. LOA_{HMT} likewise advances to 2: continuous teleoperation of the MDV is assumed in this illustrative case to require near-continuous hands-on control, to impose high operator workload, and to confine the operator to a single asset (span of control ≤ 1), placing LOA_{HMT} at the teleoperation level of the operational definition in Section 4.1. Meanwhile, LOA_S, LOA_M, and LOA_{C2} remain at LOA 1.

A2 reflects the initial MSH-II configuration with an integrated combat system, autonomous search vehicles, and semiautonomous MDV-IIs. LOA_S advances to 3 through improved autonomous search capability, LOA_M to 3 through automated path planning and replanning, LOA_F to 3 through operator-supervised semiautonomous neutralization, LOA_{C2} to 3 through combat-system-mediated coordination, and LOA_{HMT} to 3 as the operator shifts from direct control to supervision.

A3 represents the advanced MSH-II concept with networked UUV/USV search assets and operator-approved autonomous neutralization. LOA_S, LOA_M, LOA_F, and LOA_{C2} advance to 4, whereas LOA_{HMT} remains at 3: the A3 case assumes that the operator supervises and approves autonomous neutralization rather than controlling it directly, with low hands-on time, moderate workload, and a span of control greater than one, placing LOA_{HMT} at the supervisory level. Thus, the case stops short of unsupervised operation.

Table 12. Function-Specific LOA Values and Corresponding Performance Coefficients by Alternative.

Alt.	LOA_S	LOA_M	LOA_F	LOA_{C2}	LOA_{HMT}
A0	1	1 (1.00)	1 (1.00)	1 (0.65)	1 (1.00)
A1	1	1 (1.00)	2 (0.70)	1 (0.65)	2 (0.40)
A2	3	3 (1.07)	3 (1.20)	3 (0.88)	3 (0.70)
A3	4	4 (1.13)	4 (1.60)	4 (1.00)	3 (0.70)

5.4. Results

The inputs in Tables 9–12 are notional and serve only to illustrate how the framework operates; the resulting magnitudes are not performance estimates for any specific system. The value of the case lies in showing how function-specific autonomy configurations propagate through the LOA–MOP–MOE structure, not in the particular numbers obtained.

Applying Equations (7)–(12) based on inputs from Tables 9–12 and the stage-specific mission-unit assumptions yields the functional MOPs, mission-process MOPs, and mission-level MOE reported in Tables 13–15. The required quality thresholds are $P_D^{req} = 0.7$, $P_I^{req} = 0.65$, and $P_N^{req} = 0.75$, and the required route-opening time is $T_{req} = 12$ h.

Table 13. Functional MOPs by Alternative.

Alt.	V_{eff}	D_{eff}	N_{eff}^{search}	$N_{eff}^{neutralize}$
A0	5.00	1.000	1.65	1.65
A1	5.00	0.770	1.65	1.26
A2	5.89	1.452	2.85	1.62
A3	6.21	2.130	3.10	1.70

Table 14. Mission-Process MOPs by Alternative.

Alt.	T_{search}	$T_{neutralize}$	T_{MCM}
A0	4.97	21.65	26.62
A1	4.97	33.91	38.88
A2	2.38	11.63	14.01
A3	1.98	6.58	8.56

Table 15. Mission-Level MOE: Timely Route-Opening Capability.

Alt.	T_{MCM}	$T_{req}/T_{MCM,j}$	I_Q	MOE_{Open}	Interpretation
A0	26.62	0.451	0	0.000	Time and quality not met
A1	38.88	0.309	0	0.000	Time and quality not met
A2	14.01	0.857	1	0.857	Quality met; time partially short
A3	8.56	1.401	1	1.000	Time and quality both met

Under these notional values, the computation reproduces a specific pattern. Because A1 raises LOA_F and LOA_{HMT} to the teleoperation level while retaining A0's manned search stage, the assumed coefficients η_F and η_{HMT} fall below their manual baseline, so the neutralization-stage MOPs and T_{MCM} are larger at A1 than at A0 (Table 14). From A2

onward, higher cooperative and supervisory autonomy lowers T_{MCM}, and only A3 satisfies both the quality gate and the route-opening time, giving $MOE_{Open} = 1.000$ (Table 15). The interpretation of this pattern is taken up in the Discussion.

Three results deserve emphasis. First, the partial-autonomy trap appears quantitatively: introducing remotely controlled MDVs at A1 lengthens the neutralization time from 21.65 h to 33.91 h (+57%) and the total mission time from 26.62 h to 38.88 h (+46%) relative to the manned baseline, despite A1's superior neutralization equipment ($X_N = 1.1$, $P_N = 0.76$). Second, the degradation is localized: the search time is identical for A0 and A1 (4.97 h), confirming that the bottleneck lies entirely in the neutralization stage where LOA_F and LOA_{HMT} enter the teleoperation level. Third, at the mission level, only A3 satisfies both the probabilistic quality requirements and the required route-opening time ($MOE_{Open} = 1.000$); A2 satisfies the quality requirements but only partially meets the time requirement ($MOE_{Open} = 0.857$), whereas A0 and A1 fail the quality requirements ($MOE_{Open} = 0$). It should be noted that the partial-autonomy trap is observed at the mission-process MOP level rather than as a decrease in the final MOE_{Open}: because both A0 and A1 fail the quality gate, their final MOE values are identically zero, and it is the layered structure that preserves the 46.1% increase in mission-completion time at A1, which a single final score would mask.

5.5. Comparison with a Conventional Scalar-LOA Analysis

To isolate the value added by the function-specific treatment, a conventional scalar-LOA benchmark was constructed as an ablation of the proposed model. Each configuration is assigned a single system-wide autonomy level by its platform-level naming (A0 manned, level 1; A1 remote-controlled, level 2; A2 semiautonomous, level 3; A3 autonomous, level 4). In the benchmark, this single level replaces the function-specific vector: every LOA-dependent coefficient of Table 10 is evaluated at the common system-wide level. In addition, the two human–machine interaction elements—absent from the effectiveness studies reviewed in Section 3—are removed: the HMT coefficient ($\eta_{HMT} \equiv 1$; no concept of operator workload or supervisory burden) and the teleoperation penalty ($\eta_F(2) = 1$ in place of the LOA-2 valley). All other equations and inputs are identical to the proposed analysis; the technical benefits of autonomy (η_M, η_F at levels ≥ 3, η_{C2}) and the equipment gains (X, P) are retained.

The two analyses coincide at A0 (26.62 h) and diverge exactly where the removed elements act (Table 16). At A1, the benchmark predicts a 13.3% improvement (23.09 h)—the natural conclusion when the human cost of continuous teleoperation has no representation—whereas the function-specific analysis reveals a 46.1% degradation (38.88 h). At A2 and A3 the gap is attributable to the HMT axis alone: the benchmark reports 11.93 h and 7.20 h vs. 14.01 h and 8.56 h, and certifies the route-opening requirement already at A2 ($MOE_{Open} = 1.000$ vs. 0.857). Moreover, practice often does not distinguish the intermediate configurations at all and compares only the baseline and the target state; in such an A0-vs.-A3 comparison the benchmark reports a 73.0% improvement while the transition stages remain outside the analytical scope altogether.

Table 16. Conventional Scalar-LOA Analysis (ablation: no HMT coefficient, no LOA-2 valley) vs. the Staged Function-Specific Analysis (T_{MCM} in hours; MOE_{Open} in parentheses).

Alt.	Nominal System-Wide LOA	Conventional Scalar Analysis	LOA Vector	Function-Specific Analysis	Note
A0	1	26.62 (0.000)	$(1,1,1,1,1)$	26.62 (0.000)	Coincide: the removed elements are inactive when fully manned
A1	2	23.09 (0.000)	$(1,1,2,1,2)$	38.88 (0.000)	Opposite conclusions: −13.3% predicted vs. +46.1%—the trap is invisible
A2	3	11.93 (1.000)	$(3,3,3,3,3)$	14.01 (0.857)	Gap due to the HMT axis alone; requirement certified prematurely
A3	4	7.20 (1.000)	$(4,4,4,4,3)$	8.56 (1.000)	Gap due to the HMT axis alone; retained operator approval not expressible

The comparison shows what the two removed elements provide. Without the function-specific structure the bottleneck cannot be found—the scalar label attaches to the platform, not to a function—and without the interaction elements, the transition-stage degradation is replaced by a predicted gain. The staged function-specific analysis corrects the direction of the A1 prediction, localizes the degradation to teleoperated neutralization ($LOA_F = 2$) combined with operator-intensive HMT ($LOA_{HMT} = 2$), and expresses the target-state constraint of retained operator approval ($LOA_{HMT} = 3$) that scalarization cannot represent.

6. Discussion

The case is meant to show not a particular result but the kind of analysis the proposed framework makes possible under function-specific, interaction-aware modeling. Three capabilities are worth drawing out, each anchored in one of the results highlighted in Section 5.4.

First, the framework can represent non-monotonic effects of autonomy adoption. Because every function carries its own LOA and its own performance coefficient, a configuration in which partial autonomy lowers an interaction coefficient—in the case, the assumed teleoperation values of η_F and η_{HMT} at A1—can produce a mission-process MOP worse than the manual baseline even when intrinsic equipment efficiency improves. An analysis that lacks the human–machine interaction elements replaces such a stage with a predicted gain, and a single-scalar representation cannot preserve the uneven maturation of the individual functions or the diagnostic attribution of the resulting degradation; either simplification may therefore overstate or misattribute the gains of unmanned-asset introduction.

Second, the framework can localize interaction bottlenecks to particular functions. By keeping intrinsic equipment gains (X_N) separate from LOA-dependent coefficient changes (η_F, η_{HMT}) within the layered structure, it can attribute a degraded mission-process MOP to the conjunction of specific function-level conditions—in the case, LOA_F and LOA_{HMT} at the teleoperation level—rather than to the platform as a whole. The same separation yields a prescription: it indicates which functions would need to mature jointly, rather than which equipment would need to be added, to turn partial autonomy into a net gain. A single-scalar representation collapses this separation and would hide such a bottleneck.

Third, the framework preserves traceability from function-specific LOA, through functional and mission-process MOPs, to the mission-level MOE. The quality gate in Equation (5) assigns a timeliness score only when the probabilistic mission requirements are met, so the MOE cannot be reduced to a single time-saving figure and retains the

operational meaning of route opening as both timely and valid. This traceability is what would let the framework be used diagnostically once notional inputs are replaced by values grounded in test data, simulation, or structured expert elicitation.

These capabilities also clarify how the framework relates to the human–machine interaction literature on automated driving. That literature has largely characterized partial-autonomy effects at the level of the individual operator—degraded situation awareness, slower takeover responses, and trust miscalibration measured during control transitions [1–3]. What the present framework adds is a route from those operator-level interaction factors to a mission-level outcome: the same influences that driving studies treat as dependent variables enter here as the HMT coefficient η_{HMT}, which then propagates through the functional and mission-process MOPs to the MOE. In effect, the framework offers one way to ask what an out-of-the-loop or handover penalty would mean not for a single driver but for the effectiveness of a multi-asset operation. The same flexibility in the function decomposition makes the converse mapping natural: an automated-driving task would instantiate the vector without a fire or neutralization axis, retaining perception (LOA_S), maneuver and path control (LOA_M), any vehicle-to-vehicle or infrastructure coordination (LOA_{C2}), and—centrally—the driver–system teaming axis (LOA_{HMT}) that carries takeover and handover. In that setting, the framework would describe the same partial-autonomy effect that motivates this study, now expressed against a driving-appropriate MOE rather than route opening.

Beyond this structural correspondence, the qualitative pattern produced by the case is consistent with, rather than an empirical reproduction of, findings in the human–automation literature. In A1, the modeled teleoperation penalty produces mission-process performance below the manual baseline—paralleling the problematic transition zone identified for partial automation [35] and the documented costs of continuous teleoperation [29,31]—whereas the reduced operator dependence assumed from A2 onward supports recovery, consistent with the workload relief reported for higher-autonomy human–autonomy teams [27]. Notably, the two elements removed in the ablation benchmark of Section 5.5 are precisely the two elements this literature establishes as real; their removal reproduces the optimistic conclusions of the conventional stance. This consistency provides structural motivation for the coefficient pattern, but it should not be interpreted as empirical validation: workload and situation awareness were not independently measured in the present case, and the coefficient magnitudes remain notional.

The comparison in Section 5.5 also allows the prior effectiveness studies reviewed in Section 3 to be positioned rather than merely criticized. The conventional scalar benchmark should be read as a controlled abstraction of two tendencies identified in the reviewed studies—endpoint-focused evaluation and the omission or uniform treatment of autonomy [20–24]—rather than as a representation of any single prior study. Applied to the present scenario, that stance predicts a 13.3% improvement at the very stage where the interaction-aware analysis reveals a 46.1% degradation, certifies the mission requirement one stage early, and—in the endpoint-only form common in practice—reports a 73.0% improvement while the transition stages remain outside the analytical scope. The implication is not that prior end-state conclusions are wrong, but that they answer a different question—what the mature system delivers—whereas acquisition phasing, interim doctrine, and operator training depend on the question the staged analysis answers: what happens on the way there. In the same vein, the function-specific vector gives quantitative form to earlier qualitative observations, including the finding that a single LOA is insufficient for robot-swarm operations [6] and the observation that raising the LOA transforms, rather than removes, the operator's role [17].

Two broader points follow. First, the partial-autonomy trap is not specific to minehunting or to the sea. The same situation can occur in any system where a human hands part of the control to an unmanned vehicle, whether on the road, in the air, or at sea. Methods that describe autonomy with a single number, or that assume a system always runs at its highest autonomy level, tend to miss this kind of effect. Second, the value of the proposed framework is that it keeps such an effect visible and tied to the function that causes it, instead of letting it disappear into an overall average. This is the main way it differs from effectiveness analyses that treat autonomy as fixed or always high. Confirming whether a real system actually shows this dip, and how large it is, would require coefficients obtained through a rigorous evaluation process—combining engineering analysis, structured expert elicitation, and modeling and simulation to derive reliable input data—and is left to future work.

7. Conclusions

This study proposed a function-specific, five-layer framework for analyzing the effectiveness of MUM-T systems and demonstrated its analytical utility through a minehunter-centered MCM case. The framework rejects representing autonomy as a single scalar for the entire system. Instead, it defines autonomy as a function-specific vector across mission functions. LOA-dependent performance coefficients translate function-specific LOA into measurable changes in functional and mission-process MOPs. The five-layer structure comprises intrinsic technical inputs, LOA-dependent performance coefficients, functional MOPs, mission-process MOPs, and the mission-level MOE. This structure provides an explicit pathway through which autonomy operates as an operational variable shaping mission effectiveness.

As elaborated in the Discussion, the framework treats autonomy as an operational, interaction-dependent variable rather than a fixed attribute. Its central contribution is to make the human–machine interaction visible within effectiveness analysis: by carrying autonomy as a function-specific vector, it can represent non-monotonic effects of autonomy adoption, localize interaction bottlenecks—such as the partial-autonomy trap—to the functions responsible, and preserve traceability from function-specific LOA to the mission-level MOE. In this way, the study extends human–machine interaction analysis, established for automated driving, to manned–unmanned vehicle teaming. Quantitatively, under the notional inputs, the framework reproduces the partial-autonomy trap: introducing remotely controlled unmanned assets at A1 increases total mission time from 26.6 h (A0) to 38.9 h—a degradation relative to the manned baseline driven by the teleoperation penalty in η_F and η_{HMT}—before cooperative and supervisory autonomy reduce it to 14.0 h (A2) and 8.6 h (A3); correspondingly, the timely route-opening MOE is 0 for A0 and A1, 0.857 for A2, and 1.000 only for A3.

The MCM case should be interpreted as a methodological demonstration rather than as a performance forecast for any specific system. The notional inputs in Tables 9, 10 and 12, including platform-generation efficiency coefficients, probabilistic performance values, and step functions of performance coefficients, are specified to illustrate the mechanics of the framework. When these inputs are replaced by values grounded in expert judgment and technical evidence, the same five-layer structure can support analytically meaningful diagnoses of functional bottlenecks, coordinated autonomy advancement, and staged MUM-T configurations. Thus, the A3 result should be interpreted as an illustrative target-state configuration under the specified operator-approval concept, rather than as an upper bound or a forecast for any specific future minehunter configuration.

Several issues remain to be resolved for future research. A limitation of the present study is that the LOA-dependent coefficients and the probabilistic inputs were specified

notionally rather than measured; the case therefore demonstrates the structure of the framework, not calibrated performance. For application to a specific system, these inputs can be acquired through a staged path. First, the function-specific LOA levels are assigned by mapping documented system capabilities—design documentation, integration-test records, and observed behavior in operational testing—onto the established autonomy taxonomies reviewed in Section 2 (e.g., management-by-consent vs. management-by-exception for LOA_F and LOA_{C2}, or automated path-planning and replanning capability for LOA_M), while LOA_{HMT} is assigned from the three observable proxies defined in Section 4.1. Second, the framework is instantiated for the mission at hand: the function decomposition, the operationally significant MOPs, and the mission stages at which each coefficient applies are determined through engineering analysis supported by structured expert methods, such as Delphi consensus and the analytic hierarchy process (AHP) for factor prioritization. Third, once the instantiation is fixed, parameter values are estimated: the intrinsic and probabilistic inputs (X, P_D, P_I, P_N) from equipment test data and modeling and simulation; the interaction coefficients, including η_{HMT} and its LOA-2 penalty, from measurement of human-factors—operator workload (e.g., NASA-TLX [28]), situation-awareness probes, and takeover or teleoperation performance trials [29,34,35]; and, where direct measurement is infeasible, from structured expert elicitation such as Delphi-based estimation. Finally, the calibrated model is validated and refined in a closed loop: the manned baseline configuration (A0) serves, where available, as a validation anchor against operational records from past MCM exercises or comparable manned operations, and sensitivity analysis would identify the parameters that dominate mission-level outcomes—the present illustrative calculation suggests that η_F and η_{HMT} warrant particular attention—so that calibration effort can be concentrated where it matters most. This staged path makes each class of input acquirable in practice without altering the five-layer structure of the framework. The framework can also be extended to other MUM-T mission domains, such as surveillance and reconnaissance, maneuver and breaching, fire support, and security operations by tailoring the function decomposition and MOP–MOE structure to each mission process. Finally, the discrete time index t, defined herein as an upgrade-and-learning stage, can be linked to acquisition milestones and capability roadmaps, allowing the framework to support acquisition planning and the sequencing of autonomy maturation. In addition, the present case instantiates a single mission-level MOE that combines two of the MOE categories of the ME framework (Table 5): time, through the timeliness ratio, and mission satisfaction, through the probabilistic quality gate. The broader MOE set of the ME Guide—losses, expenditures, readiness, and return on investment—was not instantiated, and this choice bounds what the analysis can say about autonomy: the A1 trap and the A3 result are statements about timely route opening. Conceptually, each of the remaining categories would attach to the same function-specific LOA vector through its own pathway and would reflect different facets of autonomy and human–machine teaming. A losses MOE would have to distinguish the assets exposed inside the minefield, which differ across the alternatives (EOD personnel at A0; unmanned vehicles from A1 onward, Table 7), and would therefore weigh personnel risk and equipment attrition differently across the stages. An expenditures MOE would track consumable neutralization assets and the reprocessing driven by identification and neutralization failures—the $1/(P_I P_N)$ factor of Equation (9)—which autonomy alters through both the probabilistic inputs and the interaction coefficients. A readiness MOE would reflect the maintenance and availability of unmanned vehicles, communication links, and autonomy software—the RAM-C perspective recognized in both weapon-system ME and automated-driving safety engineering [26]. A return-on-investment MOE would relate these outcomes to the acquisition and operating costs of the staged configurations. Because the five-layer structure ties any mission-level MOE to the same function-specific

LOA vector through its own coefficients and pathways, such a multi-MOE evaluation is a direct extension of the framework; carrying it out is left to future work.

Author Contributions: Conceptualization, H.P.; methodology, G.L.; validation, J.-w.W.; investigation, H.Y.; writing—original draft preparation, G.L.; writing—review and editing, J.-w.W.; visualization, H.Y.; project administration, H.P. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by Kumoh National Institute of Technology (2024–2026) and by the Regional Innovation System & Education RISE-Specialized Industry Scale-up program through the Gyeongbuk RISE Center, funded by the Ministry of Education (MOE) and the Gyeongsangbuk-do, Republic of Korea (2026-rise-15-105).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The original contributions presented in this study are included in the article; further inquiries can be directed to the corresponding author.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Park, K. From Overtrust to Distrust: A Simulation Study on Driver Trust Calibration in Conditional Automated Driving. *Appl. Sci.* **2025**, *15*, 11342. [CrossRef]
2. Gruden, T.; Jakus, G. Determining Key Parameters with Data-Assisted Analysis of Conditionally Automated Driving. *Appl. Sci.* **2023**, *13*, 6649. [CrossRef]
3. Portron, A.; Perrotte, G.; Ollier, G.; Bougard, C.; Bourdin, C.; Vercher, J.-L. Getting Back in the Loop: Does Autonomous Driving Duration Affect Driver's Takeover Performance? *Heliyon* **2024**, *10*, e24112. [CrossRef] [PubMed]
4. AAD Knowledge Transfer Network. *Autonomous Systems: Opportunities and Challenges for the UK*; Aerospace, Aviation & Defence Knowledge Transfer Network: London, UK, 2012.
5. National Research Council. *Autonomous Vehicles in Support of Naval Operations*; The National Academies Press: Washington, DC, USA, 2005.
6. Walker, P.; Nunnally, S.; Lewis, M.; Chakraborty, N.; Sycara, K. Levels of Automation for Human Influence of Robot Swarms. In *Proceedings of the Human Factors and Ergonomics Society 57th Annual Meeting (HFES 2013), San Diego, CA, USA, 30 September–4 October 2013*; SAGE Publications: Los Angeles, CA, USA, 2013; Volume 57, pp. 429–433.
7. Huang, H.-M.; Messina, E.; Albus, J. *Autonomy Levels for Unmanned Systems (ALFUS) Framework*; National Institute of Standards and Technology: Gaithersburg, MD, USA, 2007.
8. Sheridan, T.B. *Telerobotics, Automation, and Human Supervisory Control*; MIT Press: Cambridge, MA, USA, 1992.
9. Parasuraman, R.; Sheridan, T.B.; Wickens, C.D. A Model for Types and Levels of Human Interaction with Automation. *IEEE Trans. Syst. Man Cybern. A Syst. Hum.* **2000**, *30*, 286–297. [CrossRef] [PubMed]
10. *SAE Standard J3016_202104*; Taxonomy and Definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles. SAE International: Warrendale, PA, USA, 2021.
11. Hasslacher, B.; Tilden, M.W. Living Machines. *Robot. Auton. Syst.* **1995**, *15*, 143–169. [CrossRef]
12. NATO Industrial Advisory Group (NIAG/SG-75). *Pre-Feasibility Study on UAV Autonomous Operations*; NATO: Brussels, Belgium, 2004.
13. Suresh, M.; Ghose, D. Role of Information and Communication in Redefining Unmanned Aerial Vehicle Autonomous Control Levels. *J. Aerosp. Eng.* **2009**, *22*, 85–92.
14. Li, J.; Zhao, Y.; Wang, X. A New Hierarchical Model and Evaluation Method for Autonomy Levels of Unmanned Platforms. *Res. J. Appl. Sci. Eng. Technol.* **2012**, *4*, 1343–1350.
15. Wang, Y.; Liu, J. Evaluation Methods for the Autonomy of Unmanned Systems. *Chin. Sci. Bull.* **2012**, *57*, 1740–1746. [CrossRef]
16. Young, L.A.; Yetter, J.A.; Guynn, M.D. System Analysis Applied to Autonomy: Application to High-Altitude Long-Endurance Remotely Operated Aircraft. In *Proceedings of the AIAA Infotech@Aerospace Conference, Arlington, VA, USA, 26–29 September 2005*; AIAA Paper 2005-7103; American Institute of Aeronautics and Astronautics: Reston, VA, USA, 2005.
17. Moray, N.; Rodriguez, M.; Clegg, D. Levels of Automation in Process Control. *Int. J. Cogn. Ergon.* **2000**, *4*, 51–68.
18. Frost, S.A.; Goebel, K.F.; Celaya, J.R. *A Briefing on Metrics and Risks for Autonomous Decision-Making in Aerospace Applications*; NASA Technical Report; NASA: Moffett Field, CA, USA, 2012.

19. Department of Defense. *Mission Engineering Guide*; Office of the Deputy Director for Engineering, OUSD(R&E): Washington, DC, USA, 2020.
20. Kim, B.; Lee, C. Combat Effectiveness Analysis of Manned–Unmanned Teaming System Using Agent-Based Modeling. *J. Korean Inst. Def. Technol.* **2022**, *4*, 10–18. (In Korean) [CrossRef]
21. Park, S.; Min, S.; Ha, Y. Analysis of the Combat Effectiveness of the Unmanned Surface Vehicle. *J. KNST* **2022**, *5*, 134–142. (In Korean) [CrossRef]
22. Yoon, Y.; Yoo, D.; Kim, H. Battalion-Scale Combat Effectiveness Study of Army TIGER Improvement Using AWAM. *J. Korea Acad.-Ind. Coop. Soc.* **2024**, *25*, 413–421. (In Korean) [CrossRef]
23. Lee, J.; Pyun, J.; Kim, C. A Study of MOE Establishment for Improving the Credibility of UGV Effectiveness Analysis. *J. Appl. Reliab.* **2014**, *14*, 197–202. (In Korean)
24. Shin, S.; Choi, B.; Oh, C.; Cho, H. A Case Study on Analysis Methodology of Costal Defence Weapon System. *J. Korea Inst. Mil. Sci. Technol.* **2019**, *22*, 124–134. (In Korean) [CrossRef]
25. Clough, B.T. Metrics, Schmetrics! How the Heck Do You Determine a UAV's Autonomy Anyway? In *Proceedings of the Performance Metrics for Intelligent Systems Workshop (PerMIS), Gaithersburg, MD, USA, 13–15 August 2002*; National Institute of Standards and Technology: Gaithersburg, MD, USA, 2002.
26. *ISO Standard 26262*; Road Vehicles—Functional Safety. International Organization for Standardization (ISO): Geneva, Switzerland, 2018.
27. O'Neill, T.; McNeese, N.; Barron, A.; Schelble, B. Human–Autonomy Teaming: A Review and Analysis of the Empirical Literature. *Hum. Factors* **2022**, *64*, 904–938. [PubMed]
28. Hart, S.G.; Staveland, L.E. Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research. In *Human Mental Workload*; Hancock, P.A., Meshkati, N., Eds.; North-Holland: Amsterdam, The Netherlands, 1988; pp. 139–183.
29. Rea, D.J.; Seo, S.H. Still Not Solved: A Call for Renewed Focus on User-Centered Teleoperation Interfaces. *Front. Robot. AI* **2022**, *9*, 704225. [CrossRef] [PubMed]
30. Jung, M.-a.; Lee, H.-s. A Study on the Development Strategy of the ROK Navy's Mine Countermeasure Operations: Focusing on the Analysis of Major Nations' MCM Development Trends. *J. Marit. Secur.* **2025**, *10*, 33–64. (In Korean)
31. Nitsch, V.; Färber, B. A Meta-Analysis of the Effects of Haptic Interfaces on Task Performance with Teleoperation Systems. *IEEE Trans. Haptics* **2013**, *6*, 387–398. [CrossRef] [PubMed]
32. Onnasch, L.; Wickens, C.D.; Li, H.; Manzey, D. Human Performance Consequences of Stages and Levels of Automation: An Integrated Meta-Analysis. *Hum. Factors* **2014**, *56*, 476–488. [PubMed]
33. Endsley, M.R.; Kiris, E.O. The Out-of-the-Loop Performance Problem and Level of Control in Automation. *Hum. Factors* **1995**, *37*, 381–394. [CrossRef]
34. Tan, X.; Zhang, Y. Driver Situation Awareness for Regaining Control from Conditionally Automated Vehicles: A Systematic Review of Empirical Studies. *Hum. Factors* **2025**, *67*, 367–403. [PubMed]
35. Shahini, F.; Zahabi, M. Effects of Levels of Automation and Non-Driving Related Tasks on Driver Performance and Workload: A Review of Literature and Meta-Analysis. *Appl. Ergon.* **2022**, *104*, 103824. [CrossRef] [PubMed]