# Cognitive Load-Based Curriculum Adaption in Human-Machine Teaming Training Scenarios

**Alex McConville, Nadine Marcus**
**UNSW Sydney**
**Sydney, Australia**
**a.mcconville@unsw.edu.au**

**Gary Eves**
**CAE Australia**
**Brisbane, Australia**

**Brian Stensrud**
**CAE USA**
**Florida, USA**

**Hussein Abbass**
**UNSW Canberra**
**Canberra, Australia**

## ABSTRACT

The rapid advancement and increasing deployment of autonomous systems across various operational domains present new challenges for operator training and competency development. As these systems become more intelligent and integral to mission success, there is a growing need for training that effectively prepares operators to collaborate with them. However, developing the necessary training curricula can be time-consuming, requiring instructors to align competency requirements with different system complexities and potential scenarios the operator may experience.

This paper presents a methodology for algorithmic curriculum development driven by sources of complexity for a given system. These complexity sources are described by defining context scenarios relevant to the human-machine teaming operation. Our approach identifies complexity factors within the system that influence cognitive load, then uses them to generate an initial training curriculum that builds up complexity in a simulated environment to develop operator capabilities.

The methodology incorporates physiological biometric measures of cognitive load and self-reporting of task complexity throughout the training process. By leveraging this data, alongside the operators' performance, the curriculum adjusts environmental complexity in the simulation between scenarios to balance cognitive load, promoting efficient skill acquisition and competency development.

This adaptive curriculum approach aims to ensure that trainees are neither cognitively underloaded nor overloaded to maintain an appropriate pace of personalised training. The paper discusses the application of this methodology within a multi-agent human-machine teaming framework, demonstrating how adaptive complexity management can be used as a basis for curriculum development and training within simulated environments.

## ABOUT THE AUTHORS

**Dr. Alex McConville** is a Research Associate at UNSW Sydney, School of Computer Science and Engineering. His Ph.D was on the design, development, and operation of UAVs for atmospheric sensing. His more recent work has focused on Human-Swarm Interaction for wildfire response and cognitive load in Human-Machine Teams, with a particular focus on training.

**Associate Professor Nadine Marcus** is an Associate Professor in the School of Computer Science & Engineering, UNSW, Sydney. She Coheads the Computing and Education research group. She is a Cognitive Load Theory expert who conducts research at the intersection of humans and technology, which includes the design of multimedia technology to improve human performance and learning, and novel measures of load to inform design, create adaptive learning and support human-machine teaming.

**Dr. Gary Eves** has been developing and marketing simulation systems for over 35 years. His experience started in flight, moving to cars, VR systems, high-performance computer systems, automotive, entertainment, healthcare and now back to military flight training. His doctorate is in the use of game engines to make training solutions for technical and non-technical skills training. He is currently the Principal Technology Officer for CAE IndoPac and a CAE Learning Science Fellow.

**Dr. Brian Stensrud** is currently serving as CAE's Technical Fellow for Artificial Intelligence, providing guidance and strategy on the adoption, application, and design of AI solutions across both its Defense and Civil Aviation business units. Brian has over 20 years' experience designing, developing, and deploying AI-enabled training and decision-support systems, and has led the execution of over 30 DoD-sponsored RDT&E efforts across many of the services including ARL, ONR, AFRL, and DARPA. Brian received his Ph.D. in Artificial Intelligence from the University of Central Florida in 2005, and holds undergraduate degrees in Computer Engineering, Electrical Engineering, and Mathematics from the University of Florida.

**Professor Hussein Abbass** is a Professor in the School of Systems and Computing, UNSW Canberra. He is a Fellow of IEEE and the Founding Editor-in-Chief of the IEEE Transactions on Artificial Intelligence. His current research is in Human-Swarm Teaming, Machine Education, and Test and Evaluation of AI-enabled Autonomous Systems

# Cognitive Load-Based Curriculum Adaption in Human-Machine Teaming Training Scenarios

**Alex McConville, Nadine Marcus**
**UNSW Sydney**
**Sydney, Australia**
**a.mcconville@unsw.edu.au**

**Gary Eves**
**CAE Australia**
**Brisbane, Australia**

**Brian Stensrud**
**CAE USA**
**Florida, USA**

**Hussein Abbass**
**UNSW Canberra**
**Canberra, Australia**

## INTRODUCTION

As autonomous systems grow more capable, the nature of human-machine interaction is shifting from traditional control toward collaboration. This paradigm shift is a move from conventional human-operated systems to Human-Machine Teaming (HMT), where humans and autonomous agents, with varying degrees of independence, work together toward shared goals.

This shift requires individual operators to manage greater complexity but can increase their impact, however, it also introduces the possibility of higher cognitive workload due to increased information and decision volume (Paas et al., 2022). Managing this complexity requires the delegation of some decisions to autonomous systems, which in turn depends on the development of trust, a necessary component in team performance (Hartzler et al., 2023). Training must therefore move beyond static skill acquisition to promote adaptability, trust, and shared mental models.

Human-Machine Teaming is a natural evolution of Human Computer Interaction (HCI). Building on insights from HCI where the emphasis was on usability and interface design, HMT reframes machines as proactive teammates rather than passive tools (Seeber et al., 2020). This transition builds on long-standing findings in HCI that humans apply social norms to computers, treating them as social actors (Reeves & Nass, 1996). Recognizing this, contemporary autonomous systems are now designed to display socially intelligible behaviours such as communication through natural language and offering transparency to build trust and foster collaboration (De Visser et al., 2018).

Effective teaming relies on principles like observability, predictability, and directability (Mcdermott et al., 2017), and draws from interdisciplinary research that bridges computational models with social and cognitive sciences (Neerincx, 2007). Training systems must now equip operators with the cognitive, technical, and interaction skills required for collaboration with intelligent systems. Simulations and scenario-based training should expose users to shared autonomy, dynamic role negotiation, and effective human-AI communication. This work aims to produce a general methodology to develop a modular training approach that adapts across contexts, preparing the next generation of operators to manage increasingly complex, collaborative environments.

Many defence forces still have a ‘production’ model of education, where an instructor delivers a set curriculum in a prescribed time using resources that are rarely synthetic or virtual and are used as a tool to support the instructor's experience. In such instances of training, time is the constant and competency is the variable dependent on aptitude, ability, and engagement. Purely from a pedagogical perspective, the model should change to one where competency is the consistent outcome, recognising student variability and time as commodities, particularly in areas such as defence, where competency is paramount. This presents a challenge to organisations with limited instructors, constrained scheduling ability, and a need for high levels of performance or competency. To further compound the weakness of the ‘production’ model, there is limited familiarity in the instructor cadre of operational experience of advanced autonomous systems. Additionally, the practical curriculum in using synthetic devices needs to be informed by training needs analysis, which requires operational input that is currently limited. The curriculum also needs to be informed by training media analysis, which is lacking in both autonomous systems experience and in exploiting self-paced and adaptive learning techniques.

Two further issues limit the continued use of the production model with HMTs. Firstly, the knowledge of instructors is limited by a lack of practical experience with autonomous systems, with those who are experienced needing to provide optimal front-line availability. Secondly, the current and future student demographic has been educated in the online world, used to self-paced learning, and tuition models that more often follow a ‘flipped classroom’ mode, use

of sophisticated computer games and learning by gamification, along with AI-assisted learning science embedded in common delivery platforms.

The recruitment and training of these autonomous system users must recognise the varying prior skills and abilities of students, their engagement and motivation to remain in the training system and not be lost to costly attrition. The current production model of 'one size fits all' is inadequate, and a tailored approach to competency development is required.

To tailor learning to individual progress, adaptive training systems are increasingly being used. These systems monitor performance and cognitive load to adjust task complexity or instructional content (Shute & Towle, 2003). Adaptive training could be particularly relevant for HMTs, where the interaction between human and AI agents can vary widely depending on autonomy levels, mission phases, and the roles where levels of control may vary depending on the cognitive load.

**Cognitive Load**

Cognitive load is defined as the burden placed on the working memory during problem solving or learning (Sweller, 1988), it is a critical factor in designing effective training and operational HMTs. When developing training, the working memory, responsible for learning new concepts, is limited and can be overwhelmed if it is overloaded with information (Cowan, 2001), making cognitive load management essential for effective learning, decision-making, and performance in complex environments. Cognitive load is typically divided into three types (Sweller et al., 2011), intrinsic load, stemming from the inherent complexity of the task; extraneous load, caused by suboptimal information presentation or environmental distractions; and germane load, representing the mental resources dedicated to learning. Effective instructional and interface design aims to minimize extraneous load in order to free up cognitive resources for germane load, allowing learners to focus on understanding, organizing, and integrating new information into long-term memory.

Appropriate assessment and monitoring of cognitive load is required to accurately tailor HMT systems and behaviour to human capabilities and modulate training difficulty. Physiological indicators offer a method for real-time measurement, as cognitive effort produces a response from the Autonomic Nervous System (ANS), leading to measurable changes in the physical state. Metrics such as Galvanic Skin Response (GSR), Heart Rate Variability (HRV), and Interbeat Intervals (IBI) have been positively correlated with mental workload and can act as an indirect measure of cognitive load (Ahram & Karwowski, 2016; Nourbakhsh Nargess et al., 2013; Veltman & Gaillard, 1996). Additional indicators include ocular activity (Iarlori et al., 2024), respiratory patterns (Grassmann et al., 2017) and neurophysiological signals such as EEG (Fernandez Rojas et al., 2020) can also offer insight into real-time cognitive states.

In HMT contexts, particularly in high-risk or time-sensitive environments, managing cognitive load can enhance system performance (Lindenberg & Neerincx, 2008). Systems grounded in cognitive load theory have demonstrated success and increased efficiency in human-centered automation and supervisory control of multi-agent teams (Peters et al., 2012). Emerging machine learning methods further enable adaptive interfaces capable of detecting and mitigating operator overload in real time (Jo et al., 2023; McKendrick et al., 2019). Accounting for the operator's cognitive load can be beneficial to the safe and effective deployment of collaborative autonomy.

**Training in Human-Machine Teaming**

Training for HMTs must evolve to address technical and collaborative competencies as robotic and autonomous systems increasingly support complex missions. Current training often focuses on the operation and control of robotic platforms, requiring learners to develop familiarity with interfaces and management of a robotic platform. In domains such as aviation, defence, and surgery, simulation-based training has demonstrated benefits as a cost-effective and scalable method for skill acquisition, offering exposure to high-risk, high-complexity scenarios without actual risk of harm (Cook et al., 2011). This makes the use of simulation as a training tool an ideal option for the rapidly changing HMT environment.

Simulation is particularly valuable in contexts where real-world experience is limited or potentially hazardous. In surgical education, simulation has been linked to improved procedural outcomes and safer performance, especially

when trainees practice on virtual or robotic systems before clinical exposure (Manalo et al., 2021). These findings highlight the potential of simulation for HMT training, where understanding the behaviour, limitations, and failure modes of autonomous systems is essential and cannot always be experienced in reality. Serious games are another developing area that provides interactive digital environments designed for education and skill development. They have been successfully used in military and healthcare training to engage learners in scenario-based tasks that foster decision-making, adaptability, and collaboration (Graafland et al., 2012).

As AI systems take on more agent-like roles (agentic-AI) within teams, training for HMTs begins to overlap with established practices in human-human team training. Concepts such as shared mental models, team communication, distributed cognition, and role flexibility are being considered from the perspective of human-AI collaboration (McNeese et al., 2018). This underscores the need for an approach that blends system operation, cognitive load, social understanding, and adaptive training design.

### Vignettes and Futures Studies

The value of vignettes as structured thought experiments has been shown across domains, including engineering design (Card et al., 1991), technology foresight (Boon & Park, 2005) and defence planning (Ritchey, 2006). Their utility in these contexts demonstrates the potential of scenario-based analysis not only for envisioning possible futures but for shaping the requirements of future training programs. For HMTs, this approach provides a way to overcome the current limitation of a lack of mature human-machine team systems and how to prepare human operators for roles that will demand new ways of operating. One method, General Morphological Analysis (GMA), enables the systematic exploration of a problem space by identifying all possible configurations of a system defined by a set of key mutually exclusive dimensions (Johansen, 2018).

In applying GMA to the HMT training context, we produce a wide range of potential systems, vignette development then becomes a valuable tool to develop our understanding of these configurations. In this case, a vignette describes an example of the system being applied in a potential environment. Such vignettes offer a controlled means to prototype teaming concepts, explore trust calibration, cognitive load management, and adaptive task allocation under varied future conditions (Pincombe et al., 2013). The configurations determined from GMA and the vignettes in which they are considered can be further guided using other approaches from futures studies, such as vignette development, where multiple plausible future operational environments are constructed to contextualize system performance, shape design requirements, and anticipate emerging challenges (Bishop et al., 2007).

## METHODOLOGY OVERVIEW

Conventional approaches to instructional development often rely on task needs analyses grounded in known job functions, working with well-defined system specifications. These methods assume clearly defined roles and static operational contexts, allowing for the identification of required skills and curriculum development (Mulder, 2012). However, in rapidly evolving domains where future conditions remain uncertain, such static methodologies may not be sufficient. To address this, we developed an alternative approach designed to identify the complexities and underlying skill requirements associated with potential system configurations. This approach allows us to manage the dynamic nature of emerging HMTs and support the development of training solutions that remain relevant and effective across a range of possible futures.

This work proposes a methodology to develop training schemes for HMTs, the process for which is shown in Figure 1. The approach has been considered from a system-agnostic perspective and then illustrated using an example of controlling a swarm of autonomous vehicles in a shepherding context. To determine the capabilities and training requirements of a human operator in a human-machine teaming, we initially considered the range of possible systems and reduced this to a single high-level system description. This system description is then used to define possible environments in which it could be implemented and determine the skills the operator must possess. From these skills, we define the training framework to allow operators to build the necessary competence.

The output of this methodology is a curriculum and the description or development of a training environment capable of the procedural generation of training scenarios to build the required capabilities in the operators while reducing instructor workload. The curriculum developed can be adapted based on operator capability, allowing us to move towards an individualised competency-based training approach.

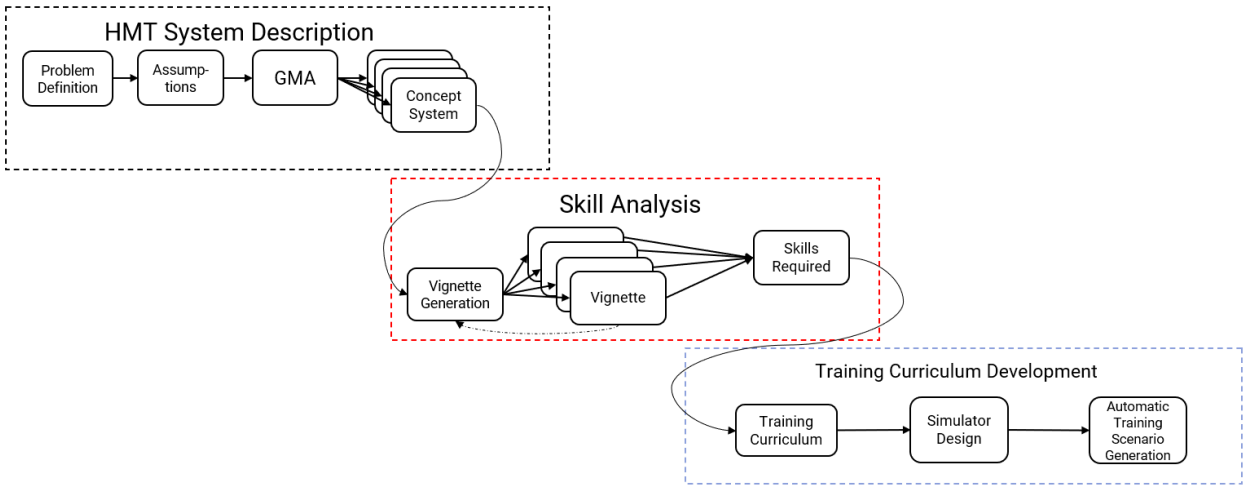


**Figure 1: Flow diagram describing the proposed training curriculum development procedure for Human-Machine Teaming**

### Human-Machine Team System Description

The initial stages of the process focus on developing a more detailed system description. The range of potential human-machine team systems is incredibly broad, and the training requirements for operators will vary significantly based on the system. In this stage, several dimensions for HMTs that are mutually exclusive are described and used to develop a range of potential systems through a systematic and traceable process.

This range is then trimmed to a more manageable number based on what is possible and logical, and clustered based on their similarity, defining potential types of system to move forward with. This is based on a GMA as discussed previously, which uses a cross-consistency matrix to remove impossible or illogical combinations of dimensions within our system descriptions. Within each type of system that has been described, there can be variations, and our description should be used as a guide for the vignette development in the next steps of the process. If the system is already under development or has been developed, then this stage can be skipped, or the focus can be on the known system capabilities.

### Skill and Complexity Analysis

Using the system description, we develop vignettes to provide additional context to the system, specific mission profiles, operational environments, team structures, system roles, and control approaches may be considered. By considering the system in context, we can understand what an operator in this type of HMT might experience and develop an understanding of the training needs.

Once a vignette has been defined, we determine the sources of complexity within that environment. We consider sources of complexity to be aspects of the environment that make achieving the goal challenging, for example, navigating a remote ground vehicle through an environment is made more challenging by the inclusion of hazardous terrain. Therefore, a source of complexity in this vignette may be the inclusion of obstacles. This process should be approached iteratively and can be considered for several variations of the system, improving the understanding of the general complexities of operating as part of this type of team. By considering the HMT in a range of contexts, we can determine the skills that underpin successful operation. These distinct skills can then be related to the sources of complexity that impact the level of skill required to manage a scenario effectively.

### Curriculum Development and Adaption

With an understanding of the skills, complexity, and a broader understanding of how a system of this type may be used from the vignettes, we can consider the training curriculum development and simulator design. This curriculum and simulator are designed with automatic scenario generation as the basis for training, where a scenario describes a

particular instance of this system in an operational environment or training case. By coupling the skills and complexity sources, we can define the necessary complexity combinations that can be used to develop the required competencies.

In this approach, we describe each complexity as a dimension of the scenario space, each scenario can then be described by its position in this space. This allows us to develop a route through the scenario space using path planning approaches such as Dijkstra's algorithm. The initial curriculum developed this way is described as a route through the scenario space, providing the skeleton for our operator training, ensuring the training progresses in a logical manner, and the required competencies are developed.

By measuring task performance and the cognitive load of trainees in a manner appropriate for the system, we can adjust the complexity of the generated scenarios for the individual. As the training environment allows for automatic scenario generation based on the complexity description from the curriculum, we can produce an individual training program based on the trainee's capabilities. By personalising the pace of training to cater to the individual, we can maintain engagement and ensure enough cognitive resources are available to allow learning to take place effectively.

### Training Environment Development

This process allows us to produce a modular training environment for our system and develop the competencies required to manage the determined complexity sources. We structure a training environment as four distinct components: a curriculum generator, a scenario simulator, a Human-Machine Interface (HMI), and a curriculum adaptation engine. Each component can be modified or adjusted without impacting the other systems, assuming they are compatible with the inputs and outputs from the other components. The system is designed to allow components to be reused in different contexts or modified for changing circumstances without having to develop an entirely new system. The flow through the environment described is displayed in Figure 2.

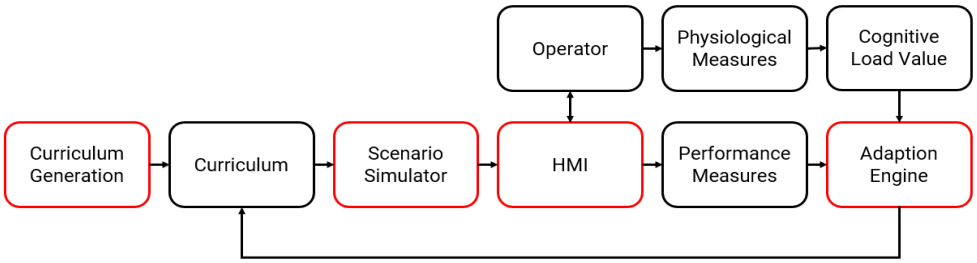


**Figure 2: The flow experienced by a trainee using the training environment**

## EXAMPLE IMPLEMENTATION

Here we provide an example to demonstrate an application of our methodology. Throughout this example, we discuss a possible approach to each step in the methodology and follow the process to the development of an example training environment. This is just one possible approach to implementing this methodology for one use case. It is meant to offer some guidance on the implementation for possible applications.

### Human Machine Team System Description

To determine a HMT description we consider the assumptions and differentiating factors within the system space. Here, we assume that there is a single human operator in the team, and the intelligence-based autonomy levels are in line with the levels described by (Sheridan & Verplanck, 1978) and are technologically achievable. The dimensions used to describe our system are described in Table 1.

The dimensions describe: the density of the agents, expressing how closely these agents collaborate on tasks, and their general physical distribution from each other. The number of hardware agents in the team ranges from a single collaborative teammate to a swarm of a significant scale. The relative location of the operator impacts the level of awareness and ability to intervene under difficult conditions. The level of automation the agents are capable of, and the position of the intelligence in the system, describe where the intelligence sits within the system, which may impact the operations based on additional conditions in the environment.

**Table 1: Human Machine Team Dimensions.**

| Density | Hardware Agents | Operator Location | Automation Level | AI Position |
|---|---|---|---|---|
| Low | 1 | Line of Sight | 3 | Agent |
| Medium | 2 – 20 | In Region | 5 | Controller |
| High | 20+ | Fully Remote | 7 | Interface |

From these dimensions, there is a total of 243 possible configurations in our HMT design space. We reduced the number by eliminating logically inconsistent or impossible combinations using a Cross-Consistency Matrix (CCM), shown in Figure 3. For example, a high number of low-autonomy agents that would overload the operator and single-agent systems labelled as high-density were excluded. After removing these configurations that include conflicting pairs, 62 valid configurations remained. These were then grouped using K-modes clustering into six distinct clusters. Configuration 1, shown in Table 2 was selected, featuring a medium-sized swarm of highly autonomous agents managed remotely.

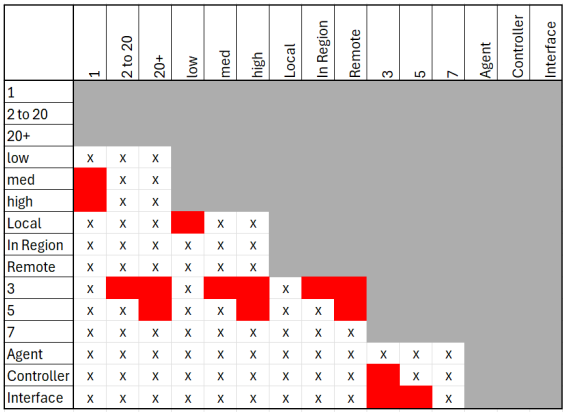


**Figure 3: Cross Consistency Matrix with red highlighting inconsistent or impossible pairings**

**Table 2: System Configuration List**

| | Density | Hardware Agents | Operator Location | Automation Level | AI Position |
|---|---|---|---|---|---|
| 1 | Med | 2 – 20 | Fully Remote | 7 | Agent |
| 2 | High | 20+ | In Region | 7 | Interface |
| 3 | Low | 1 | In Region | 5 | Controller |
| 4 | Low | 20+ | Fully Remote | 7 | Controller |
| 5 | Med | 2 – 20 | Line of Sight | 5 | Controller |
| 6 | Low | 2 – 20 | In Region | 7 | Agent |

**Contextual Vignette and Complexity Sources**

Given the description of the HMT we have previously determined, we need to then develop a contextual vignette for this HMT. Through this, we develop a deeper understanding of the concept of operation for this system, as well as the complexities in the environment.

**Example Vignette**: Consider an environment where several Uncrewed Underwater Vehicles (UUVs) have been distributed to collect samples in a region. These vehicles lack significant communication capabilities, so they can only be communicated with when within a short range. To control them, a control agent, a Uncrewed Surface Vessel (USV), must navigate the environment and collect them to guide them back to their collection point. These agents, once within communication range of another UGV, remain in communication range and can be navigated together, acting as a herd. The environment they are sampling is littered with obstacles and hazards that they must automatically avoid for safety reasons.

From this vignette, we can consider the sources of complexity in the environment that would make completing this task more challenging. They were determined to be the number of agents that need to be collected, their spatial distribution, the number of obstacles in the environment, and the scale of those obstacles. Given the simple environment we are describing, these are the four sources of complexity that drive the intrinsic load the operator would experience.

In this example, there are two primary skills we can determine are required for this task: collecting and driving. Collecting is efficiently herding the agents together to be guided to the desired location, and driving is effectively navigating the control agent and the herd to the desired location. Both skills require other subskills such as mission

and route planning, spatial awareness, and effective vehicle control, however, these two skills give us measurable metrics we can use to assess operator performance and can be directly related to the sources of complexity.

**Curriculum Development**

To structure the training curriculum, we define a scenario space based on the complexity sources determined from the vignette and represent them as a vector of complexity values: [number of agents, spatial distribution, obstacle density, and obstacle scale]. We also describe the skills or competencies that must be developed as a combination of these complexities. In this case, the skills determined are collecting and driving. The collect skill is primarily affected by the number of agents and their distribution, and the driving skill is affected by the number of agents, the number of obstacles, and their size. Each source is rated on a scale from 0, minimally present in the scenario, to 4, highest likely prevalence. This allows scenarios to be represented as vectors, for example, [2,2,0,0] indicates moderate collecting complexity due to the number of agents and their distribution with minimal environmental challenges, while [2,0,2,2] shifts focus to navigation and control skills for the driving competency.

We model the scenario space as a graph where each node is a scenario vector, and edges represent incremental changes (±1) in complexity. Using Dijkstra's algorithm, we generate an optimal learning path from a baseline scenario [0,0,0,0] to a final competency goal [4,4,4,4], incorporating key intermediate waypoints, such as [3,3,0,0] for collecting and [3,0,3,3] for driving. Additional constraints can guide the curriculum's progression. We penalize abrupt increases in scenario variance to promote gradual skill acquisition, and heuristics remove steps that decrease overall complexity. This leads to the developed curriculum described below.

**[(0, 0, 0, 0), (0, 1, 0, 0), (1, 1, 0, 0), (1, 2, 0, 0), (2, 2, 0, 0), (2, 3, 0, 0), (3, 3, 0, 0), (3, 2, 0, 1), (3, 1, 1, 1), (3, 0, 1, 2), (3, 0, 2, 2), (3, 0, 2, 3), (3, 0, 3, 3), (3, 1, 3, 3), (3, 2, 3, 3), (3, 3, 3, 3), (3, 3, 3, 4), (3, 3, 4, 4), (3, 4, 4, 4), (4, 4, 4, 4)].**

**Training Environment Overview**

The example training environment developed for this context is shown in Figure 4-1). In this environment, we see several entities displayed, some of which are using NATO joint military symbology (Beckman, 2025). In the figure, A) represents the goal location, which is the area where the autonomous agents are being guided to by the control agent. The symbol shown in B) represents the control agent, which is directly controlled by the operator and is used to herd the autonomous agents in the environment. The autonomous agents, represented at C, are guided by the actions of the control agent, they avoid obstacles and each other, and have herding behaviours when within a set distance of other autonomous agents. D) shows an environmental hazard that the autonomous agents avoid, however, the operator must direct the control agent to avoid them. Each scenario has a time limit to add additional load to the operator, this time limit is derived as a function of the overall complexity level of the scenario. The scenario ends when the autonomous agents have all reached the goal location or time runs out. The performance measures being used are the overall time to complete the task or the task completion percentage in case of failure.

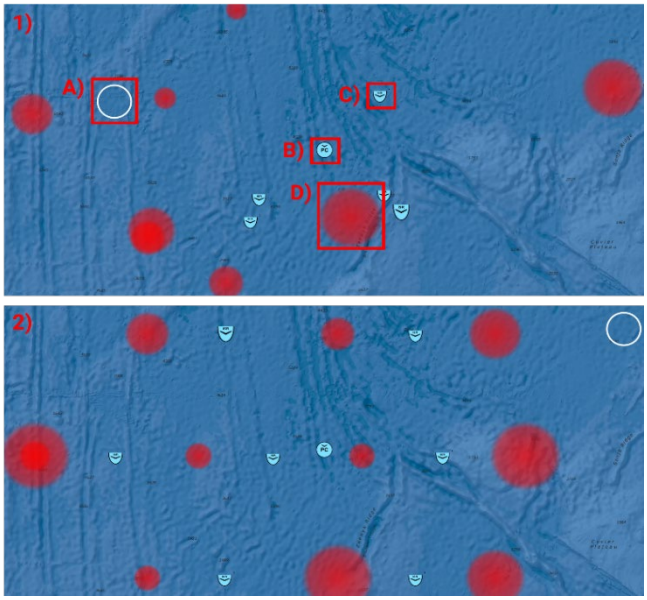


**Figure 4: Two Examples of the Training Environment with Image 1) showing an example scenario showing A) the Goal, B) the Control Agent, C) the Autonomous Agents to be herded, and D) an Environmental Hazard, and Image 2) showing an example of an Assessment Scenario**

The Human Machine Interface (HMI) for our environment is a noise-free, fully observable environment (i.e. full situational awareness display) where the operator has a clear picture of the overall environment, including the location of the goal, all obstacles, and all autonomous agents. It provides minimal external information and gives no information on the agents beyond their position in space.

Alternative HMIs would impact how the operator interacts with the environment and could be applied to describe different scenarios.

Within our environment, we have included the capacity for predefined scenarios, allowing a consistent tutorial section, ensuring all trainees experience the same information before starting the curriculum, and predefined assessment stages to ensure consistency of experience across trainees. The environment displayed in Figure 4-2) shows an assessment stage further along the curriculum with a higher level of complexity. The positions of the obstacles, agents, and goal are all predefined in the environment to allow consistency in the evaluation of trainees at this point in the curriculum.

**Curriculum Adaption**

When developing the training environment, due to the procedurally generated nature of the scenarios, no lesson outside of the tutorial or assessments is predefined, which gives us the capacity to adjust the scenarios for the individual. The initial curriculum provides a route or heading to develop the competencies required, however, as individuals grow accustomed to different complexities or struggle with the cognitive load of the environment, we can adjust the pace and specifics of the training program they experience.

The adaption engine uses Heart Rate Variability (HRV) and Galvanic Skin Response (GSR) as physiological indicators of the load experienced during this scenario. Following each scenario, the operator is then given a NASA-TLX to self-assess their cognitive load. These values are normalised and combined to produce a single cognitive load value. This value, alongside an overall task performance metric, are inputs to determine if or how the next scenario should deviate from the curriculum for this individual. In this example system, a simple model has been used to describe the approach to adaption. Depending on the regions displayed in Figure 5, where the operator's performance and cognitive load fall after each scenario, the curriculum is adapted in a pre-described way based on Table 3. As the operator progresses through the curriculum, it will adapt to meet their training needs, moving quickly through areas they excel at and slowing to develop skills in the areas requiring more development. Developing a coaching system to understand why the operators are succeeding or failing sits outside the scope of this work, but would be beneficial to implement alongside an adaptive simulation environment.

**Table 3: Adaption based on Cognitive Load and Performance state**

| Region | Curriculum adaption |
|---|---|
| 1 | Repeat scenario – explain the goal |
| 2 | Repeat scenario – explain tasks and skills |
| 3 | Step back through the curriculum to the last achievable complexity state or return to the tutorial |
| 4 | Step forward through the curriculum |
| 5 | Maintain overall complexity, adjust scenario based on which skill performance is limiting overall performance |
| 6 | Reduce complexity. Focus on maintaining complexity in the skill performance that is limiting success |
| 7 | Step forward through the curriculum |
| 8 | Half-step forward through the curriculum |
| 9 | Reduce complexity. Focus on maintaining complexity in the skill performance that is limiting success |

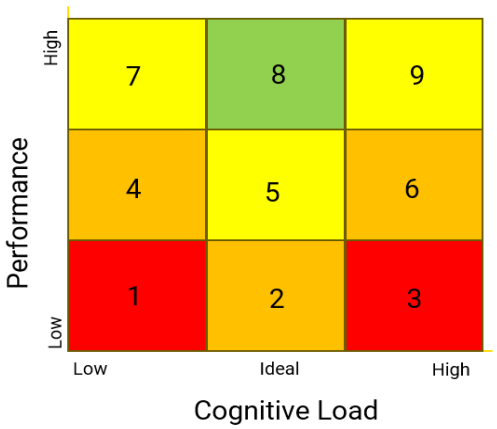


**Figure 5: Grid displaying different operator state regions based on performance and cognitive load throughout a scenario**

**DISCUSSION**

The purpose of this work is to improve preparedness for working in human-machine teaming environments, which will develop and shift over the coming decades. This work is focused on developing a way to determine and develop the general competencies required for different HMTs, even if the precise use case or mode of operation is unknown, and doing so in a flexible, adaptable, and individual way. The approach is centred on defining the overarching sources of complexity that can be considered across contexts and industries, and defining the skills required as a combination of these complexities. Describing the environment in this way provides the opportunity to build a modular training environment that can be adjusted to the organisational needs, changing system capabilities, interfaces and developing environments without requiring trainers to develop new curricula manually.

Throughout this work, we have described an approach to instructional design that does not require a specific job role, well-defined task, or concrete application. By extracting the complexities and skills from contextual vignettes, we can begin to develop our understanding of what it means to be part of these teams and infer the capabilities and tasks that will be required. This approach to describing the complexities as individual dimensions in an N-dimensional space opens up opportunities to apply path or trajectory planning approaches to curriculum design. Using these complexity dimensions as inputs to procedural simulation environments provides an easily modifiable approach to implementing this curriculum, and integrating this with cognitive load measurement of the trainee allows highly personalised training to ensure the necessary competencies are reached.

The primary limitation of this approach originates from the qualitative nature of determining the complexities in the environment. The complexities described may differ based on the perspectives of different training developers. This can have significant downstream effects and highlights the importance of getting relevant stakeholders involved in the training development process, as well as iterating on the context scenarios and considering a wide variety of potential environments, objectives, and purposes of the HMT. If the complexity sources used as the basis for curriculum and simulator design are lacking, then the overall training environment may not be fit for purpose, and the curriculum and, more significantly, the simulator may have to be updated.

While the focus of this work was on the development of adaptive computer-based training for HMTs, the principles described here could also be applied to real-world scenario-based training approaches not focused on HMTs. If the sources of complexity can be described sufficiently, a valid curriculum and scenario generation system may be able to generate useful training scenario descriptions for trainers where the simulation component takes place in a physical environment.

Additionally, by developing this modular training environment, we can provide a useful research testbed, which allows us to assess different HMIs, the effectiveness of different curricula, and a range of approaches to curriculum personalisation, all for a single simulation system. Or we could apply a previously built curriculum generation, HMIs, and adaption mechanism to a different scenario simulation environment, reducing workload for those developing and delivering the training.

Ultimately, by shifting the focus from specific roles to the underlying complexities of operating as part of a human-machine team, this methodology provides a flexible and adaptive approach to training design that can evolve alongside the technological and operational demands.

**ACKNOWLEDGEMENTS**
This research was supported as part of the Australian Defence Trailblazer RAS AI Program.

**REFERENCES**

Ahram, T. Z., & Karwowski, W. (2016). *Advances in Intelligent Systems and Computing 494 Advances in The Human Side of Service Engineering*. http://www.springer.com/series/11156

Beckman, M. (2025, June 23). *Milsymbol*. Https://Github.Com/Spatialillusions/Milsymbol.

Bishop, P., Hines, A., & Collins, T. (2007). The current state of scenario development: An overview of techniques. In *foresight* (Vol. 9, Issue 1, pp. 5–25). https://doi.org/10.1108/14636680710727516

Boon, B., & Park, Y. (2005). A systematic approach for identifying technology opportunities: Keyword-based morphology analysis. *Technological Forecasting and Social Change*, *72*(2), 145–160. https://doi.org/10.1016/j.techfore.2004.08.011

Card, S. K., Mackinlay, J. D., & Robertson, G. G. (1991). A Morphological Analysis of the Design Space of Input Devices. *ACM Transactions on Information Systems*.

Chen, D., Yoon, H. J., Wan, Z., Tech, V., Alluru, N., Lee, S. W., He, R., Woodson, C. G., Moore, T. J., Nelson, F. F., Yoon, S., Lim, H., Kim, D. D., & Cho, J.-H. (2018). *Advancing Human-Machine Teaming: Concepts, Challenges, and Applications* (Vol. 1).

Cook, D. A., Hatala, R., Brydges, R., Zendejas, B., Szostek, J. H., Wang, A. T., Erwin, P. J., Hamstra, S. J., & Corresponding, H. ). (2011). Technology-Enhanced Simulation for Health Professions Education A Systematic Review and Meta-analysis. In *JAMA* (Vol. 306, Issue 9). www.jama.com

Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences*, *24*(1), 87–114. https://doi.org/10.1017/S0140525X01003922

De Visser, E. J., Pak, R., & Shaw, T. H. (2018). From ‘automation’ to ‘autonomy’: the importance of trust repair in human–machine interaction. *Ergonomics*, *61*(10), 1409–1427. https://doi.org/10.1080/00140139.2018.1457725

Fernandez Rojas, R., Debie, E., Fidock, J., Barlow, M., Kasmarik, K., Anavatti, S., Garratt, M., & Abbass, H. (2020). Electroencephalographic Workload Indicators During Teleoperation of an Unmanned Aerial Vehicle Shepherding a Swarm of Unmanned Ground Vehicles in Contested Environments. *Frontiers in Neuroscience*, *14*. https://doi.org/10.3389/fnins.2020.00040

Graafland, M., Schraagen, J. M., & Schijven, M. P. (2012). Systematic review of serious games for medical education and surgical skills training. In *British Journal of Surgery* (Vol. 99, Issue 10, pp. 1322–1330). https://doi.org/10.1002/bjs.8819

Grassmann, M., Vlemincx, E., von Leupoldt, A., & Van den Bergh, O. (2017). Individual differences in cardiorespiratory measures of mental workload: An investigation of negative affectivity and cognitive avoidant coping in pilot candidates. *Applied Ergonomics*, *59*, 274–282. https://doi.org/10.1016/j.apergo.2016.09.006

Hartzler, B. M., Scielzo, S., Abraham, A., Wong Spencer Kohn, R., Church, F., & BethHartzler, V. (2023). Effects of Trust Calibration on Human-Machine Team Performance in Operational Environments. *Interservice/Industry Training, Simulation, and Education Conference (I/ITSEC)*.

Iarlori, S., Perpetuini, D., Tritto, M., Cardone, D., Tiberio, A., Chinthakindi, M., Filippini, C., Cavanini, L., Freddi, A., Ferracuti, F., Merla, A., & Monteriù, A. (2024). An Overview of Approaches and Methods for the Cognitive Workload Estimation in Human–Machine Interaction Scenarios through Wearables Sensors. In *BioMedInformatics* (Vol. 4, Issue 2, pp. 1155–1173). Multidisciplinary Digital Publishing Institute (MDPI). https://doi.org/10.3390/biomedinformatics4020064

Jo, W., Wang, R., Yang, B., Foti, D., Rastgaar, M., & Min, B.-C. (2023). Cognitive Load-based Affective Workload Allocation for Multi-human Multi-robot Teams. *IEEE Transactions on Human Machine Systems*, *55*(1). http://arxiv.org/abs/2303.10465

Johansen, I. (2018). Scenario modelling with morphological analysis. *Technological Forecasting and Social Change*, *126*, 116–125. https://doi.org/10.1016/j.techfore.2017.05.016

Lindenberg, J., & Neerincx, M. (2008). Situated cognitive engineering for complex task environments. In *Naturalistic Decision Making and Macrocognition*. https://www.researchgate.net/publication/287918596

Manalo, T. A., Higgins, M. I., Pettitt-Schieber, B., Pettitt, B. J., Mehta, A., & Hartsell, L. M. (2021). A Strategy for Undergraduate Medical Education in Urology During the COVID-19 Pandemic. *Journal of Surgical Education*, *78*(3), 746–750. https://doi.org/10.1016/j.jsurg.2020.09.011

Mcdermott, P. L., Walker, K. E., Dominguez, C. O., Nelson, A., & Kasdaglis, N. (2017). Quenching the Thirst for Human-Machine Teaming Guidance: Helping Military Systems Acquisition Leverage Cognitive Engineering Research. *13th International Conference on Naturalistic Decision Making 2017*.

McKendrick, R., Feest, B., Harwood, A., & Falcone, B. (2019). Theories and Methods for Labeling Cognitive Workload: Classification and Transfer Learning. *Frontiers in Human Neuroscience*, *13*. https://doi.org/10.3389/fnhum.2019.00295

McNeese, N. J., Demir, M., Cooke, N. J., & Myers, C. (2018). Teaming With a Synthetic Teammate: Insights into Human-Autonomy Teaming. *Human Factors*, *60*(2), 262–273. https://doi.org/10.1177/0018720817743223

Mulder, M. (2012). Competence-Based Education and Training. *Journal of Agricultural Education and Extension*, 305–314.

Nourbakhsh Nargess, Wang Yang, Chen Fang, & Calvo Rafael A. (2013). *Using Galvanic Skin Response for Cognitive Load Measurement in Arithmetic and Reading Tasks*.

Paas, A., Coffey, E. B. J., Beltrame, G., & St-Onge, D. (2022). *Towards evaluating the impact of swarm robotic control strategy on operators’ cognitive load*. https://www.bitcraze.io/products/crazyflie-2-1/

Peters, J. R., Srivastava, V., Taylor, G. S., Surana, A., Eckstein, M. P., & Bullo, F. (2012). Human Supervisory Control of Robotic Teams Integrating Cognitive Modeling with Engineering Design. *IEEE Control Systems Magazine*.

Pincombe, B., Blunden, S., Pincombe, A., & Dexter, P. (2013). Ascertaining a hierarchy of dimensions from time-poor experts: Linking tactical vignettes to strategic scenarios. *Technological Forecasting and Social Change*, *80*(4), 584–598. https://doi.org/10.1016/j.techfore.2012.05.001

Reeves, B., & Nass, C. (1996). *The Media Equation: How People Treat Computers, Television, and New Media Like Real People and Pla*. https://www.researchgate.net/publication/37705092

Ritchey, T. (2006). *Modeling Multi-Hazard Disaster Reduction Strategies with Computer-Aided Morphological Analysis*. www.swemorph.com]

Seeber, I., Bittner, E., Briggs, R. O., de Vreede, T., de Vreede, G. J., Elkins, A., Maier, R., Merz, A. B., Oeste-Reiß, S., Randrup, N., Schwabe, G., & Söllner, M. (2020). Machines as teammates: A research agenda on AI in team collaboration. *Information and Management*, *57*(2). https://doi.org/10.1016/j.im.2019.103174

Sheridan, T. B., & Verplanck, W. (1978). *Human and Computer Control of Undersea Teleoperators*. https://www.researchgate.net/publication/23882567

Shute, V., & Towle, B. (2003). Adaptive e-learning. *Educational Psychologist*, *38*(2), 105–114. https://doi.org/10.1207/S15326985EP3802_5

Sweller, J. (1988). Cognitive Load During Problem Solving: Effects on Learning. *Cognitive Science*, *12*(2), 257–285. https://doi.org/10.1207/s15516709cog1202_4

Sweller, J., Ayres, P., & Kalyuga, S. (2011). *Explorations in the Learning Sciences, Instructional Systems and Performance Technologies: Cognitive Load Theory*. http://www.springer.com/series/8640

Veltman, J. A., & Gaillard, A. W. K. (1996). Physiological indices of workload in a simulated flight task. In *Biological Psychology* (Vol. 42).