Proceedings of the 2023 Winter Simulation Conference
C. G. Corlu, S. R. Hunter, H. Lam, B. S. Onggo, J. Shortle, and B. Biller, eds.

EVOLVING LVC TO INCLUDE EVALUATION OF HUMAN-AI TEAMING DYNAMICS

Margaret Loper

Information and Communications Laboratory
Georgia Tech Research Institute
250 14th Street NW
Atlanta, GA 30318, USA

Valerie Sitterle

Electronic Systems Laboratory
Georgia Tech Research Institute
250 14th Street NW
Atlanta, GA 30318, USA

ABSTRACT

There are significant differences between using systems as human-controlled tools to accomplish a specific task and using systems designed to “cooperate and partner” with humans to achieve capabilities beyond either side acting alone. The live, virtual, constructive (LVC) paradigm increasingly emphasized by the DoD has wide acceptance and is congruent with how the military thinks about training, evaluation, and mission rehearsal. Consequently, it may help address these challenges. This paper aims to overview the current LVC construct, challenges associated with human-AI teaming and intentional design of these dynamics to achieve new capabilities, and the resulting need to evolve the LVC construct to improve our pursuit of understanding and evaluation that leads to effective fielding.

1 INTRODUCTION

There are significant differences between using systems as human-controlled tools to accomplish a specific task and using systems designed to “cooperate and partner” with humans to achieve capabilities beyond either side acting alone. Systems intended to team with humans are increasingly enabled by artificial intelligence (AI), with varying degrees of adaptive or “learning” potential. Recognizing there is an entire spectrum of how reasoning, learning, and adaptation in algorithmic form is manifested, there is similarly a spectrum of how teaming may be realized across the spectrum of algorithmic capabilities. Comprehensive testing of all possible states is impossible or infeasible and the evaluative intractability inhibits our abilities to understand and design these relationships for effective use. We need to understand how to be intentional across various forms of teaming, and this has critical relevance for enhancing operational capabilities.

The US Department of Defense (DoD) needs to train for and evaluate the effectiveness of humans teaming with AI-enabled systems across the high degree of potential variation to this dynamic to achieve and field new capabilities that operate as intended and are safe for use in theater. At present, however, there remains a capability gap with respect to designing human-AI teams and using training to drive innovative design and evaluation of those designs to support further training, design, and evaluation. The live, virtual, constructive (LVC) paradigm increasingly emphasized by the DoD has wide acceptance and is congruent with how the military thinks about training, evaluation, and mission rehearsal. While the LVC construct may help address these challenges, it is not yet designed for nor is there solid understanding of how we might apply it toward the problem of human-AI teaming.

This paper aims to overview the current LVC construct, challenges associated with human-AI teaming and intentional design of these dynamics to achieve new capabilities, and the resulting need to evolve the LVC construct to improve our pursuit of understanding and evaluation that leads to effective fielding.

979-8-3503-6966-3/23/$31.00 ©2023 IEEE

2 THE LVC CONSTRUCT

The LVC construct, which categorizes the way humans interact with simulations, first originated in 1991. Inspired by the Army's challenge to increase the effectiveness of training methods to meet the challenges posed by full spectrum warfare, GEN Paul F. Gorman presented a paper to the Society for Computer Simulation that argued "…most military training could be advantaged by Tactical Engagement Simulation (TES) in any or all of its three forms, Constructive, Virtual and Subsistent, and that, ideally, all three forms would be used interactively" (Gorman 1991). As documented in his Interservice/Industry Training, Simulation and Education Conference (I/ITSEC) Fellow's paper, the terms in the 1991 paper morphed over time: "subsistent" became "live" and Seamless TES became "blended training" (Gorman 2011). Later, the Defense Science Board (DSB) Task Force on Simulation Readiness & Prototyping, on which GEN Gorman was a member, solidified the concept of LVC, by asserting "Everything is simulation except combat" and classified the whole of Modeling, Simulation and Gaming as Live, Constructive, and Virtual (DSB 1992).

2.1 LVC Defined

With this background in place, a commonly used figure capturing the LVC taxonomy emerged in 2013, as shown in Figure 1 (I/ITSEC 2013). Live simulation refers to Modeling & Simulation (M&S) involving real people operating real systems (e.g., a pilot flying a jet) for a simulated mission. A virtual simulation is one that involves real people operating simulated systems (e.g., a pilot flying a simulated jet). Constructive simulations are those that involve simulated people operating simulated systems (e.g., a simulated pilot flying a simulated jet). According to the matrix in Figure 1, there is no name for simulated people operating real equipment. When the LVC taxonomy was created in 1991, there were no examples of this type of interaction. However, technology has advanced to the point where simulated humans are operating real systems. For example, the DARPA AlphaDogfight Trial was a competition where AI controlled a simulated F-16 fighter jet in aerial combat against an Air Force pilot flying in a virtual reality simulator (Goecks 2022). The winner defeated an experienced pilot in a simulated dogfight, demonstrating that an AI-based pilot may surpass human abilities. This quadrant has not been named, but is sometimes called Autonomy.

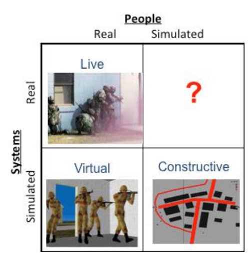

Figure 1: Categorizing simulations by the way humans interact with them.

2.2 LVC Interoperability Challenges

To explain some of the interoperability challenges of LVC simulations, we will start by describing the basic elements of a simulation: objects, events, behavior, environment, time, and data. Objects represent things such as people, vehicles, sensors or computers that are modeled in the simulation. An Event is an instantaneous occurrence that changes the state of the system; and each event has a time associated with it indicating when it occurred. Behavior represents the actions and interactions of the objects; it is the unfolding of events over time. The Environment in which those objects exhibit behavior could include land, sea, air, space, or none. Since a simulation is a method for exercising a model over time, Time is represented by a clock which is advanced as a result of the occurrence of events. Lastly, Data represents the parameters in the simulation as well as the scenario, and captures things like fidelity. To briefly explain some of the interoperability challenges associated with LVC, consider the notional scenario, represented by the center box in Figure 2, which includes aircraft, helicopters, ships, and land vehicles.

On the far left of Figure 2 is a Virtual Aircraft. A Virtual simulation is one that involves real people operating simulated systems, which means that the objects and behaviors in the simulation will include both human (pilot) and models (aircraft). The environment will be simulated (land and sea), and time will advance with wallclock (real-time) to support the human decision making of the pilot. On the bottom of Figure 2 is a Live Helicopter. A Live simulation involves real people operating real systems, which means the objects and behaviors are human (pilot) and real (helicopter). The environment is real, and time will advance with wallclock to support the human decision making of the pilot. On the right of Figure 2 is Constructive Units. A Constructive simulation involves simulated (or no) people operating simulated systems, which means the objects and behaviors are represented as models (people, ships, vehicles). The environment is simulated (land and sea) and time will advance with the simulation clock managed locally (slower or faster than real time).

Each of these LVC simulations exchange state information about their objects and behaviors (e.g., object type, time, position, velocity, weapons fire, emissions) repeatedly throughout the execution of the scenario. From this discussion, it is easy to illustrate the numerous issues associated with integrating simulations across LVC. For example, models of systems may have different mathematical representations and performance than real counterpoints, correlating real and synthetic environmental representations can introduce significant error, coordinate systems of real platforms and models have to be aligned, the aggregation of the objects (e.g., single helicopter or combination of planes) needs to be considered, wallclock and simulation clocks synchronized, causal ordering of events maintained, and still there are other issues such as network latency, security, etc. Some of these issues can be solved with standards, which is where Distributed Interactive Simulation, High Level Architecture, and Test and Training Enabling Architecture come into play. In other cases, for example semantic interoperability, standards are not enough. These known issues have been documented through numerous papers over many decades. However, they will become more complex when we introduce AI teaming with LVC simulations.

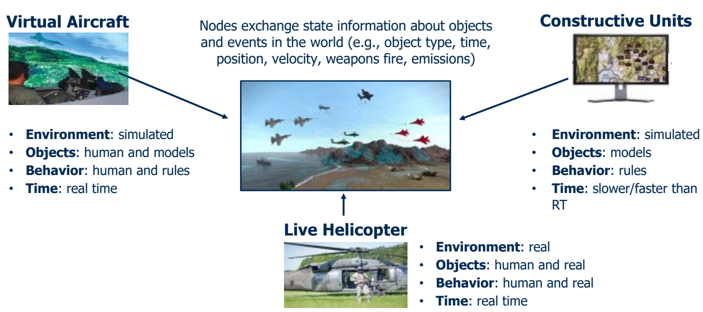

Figure 2: LVC interoperability challenges.

3 AI TEAMING

3.1 Spectrum of System Capabilities

Of paramount importance to design behavior dynamics and desired outcomes for systems including AI elements is the understanding that AI and machine learning (ML) capabilities cannot be treated monolithically. The adaptivity and autonomy an AI-enabled system demonstrates (or is allowed to exhibit) drives overall system and hybrid system behaviors. A wide range of AI and ML approaches imbue systems with varying capability and behavior profiles, and researchers characterize them in different ways. Parasuraman et al. (2000), for example, employs Sheridan's 10-point scale for levels of automation to characterize four primary automation functions: information acquisition, information analysis, decision and action selection, and action implementation. New capabilities in autonomy, enabled through AI and ML, extend beyond automation. They challenge current DoD approaches to system design, validation, and – importantly – to how we develop the confidence and knowledge to use these systems in theater. Figure 3 illustrates a high-level characterization of systems with functionality driven by AI or ML; it is heavily influenced by considerations for operational use in defense situations. ML is taken as a subset of AI, where the former uses specific algorithm types and the latter is the broader concept with highly varying approaches seeking to enable a machine or system to sense, reason, act, or adapt. General AI is considered as AI with the flexibility, resourcefulness, and self-awareness of humans, a conventional yet debated definition.

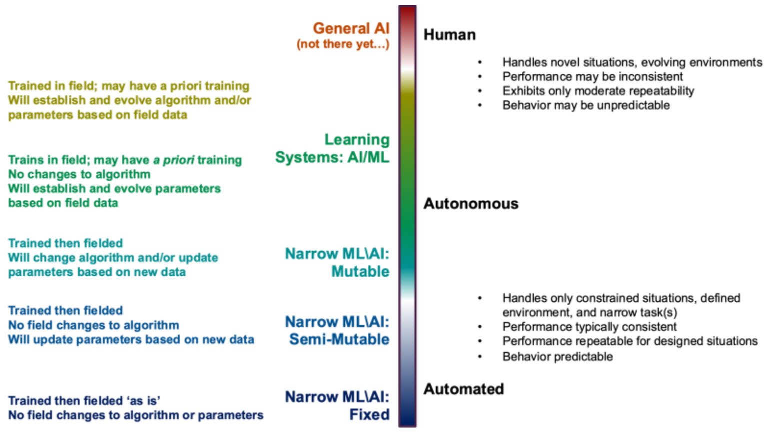

Figure 3: High-level classification spectrum of ML, AI, and learning-enabled systems for LVC consideration, adapted from (Collopy and Sitterle 2019).

Narrow ML or AI is typically task-specific and fixed in terms of its algorithm code and parameters once trained. Moreover, strictly ML-based algorithms based on pattern recognition offer no inherent causal inference. These systems exhibit the same brittleness and considerations as other automated engineering systems, and are consequently already accounted for in the existing LVC construct. The challenge occurs as we move toward ML and AI systems able to handle greater numbers of operational situations with greater degrees of autonomy. Yang et al. (2020) define algorithmic complexity as probabilistic, adaptive, evolving probabilistic, or evolving adaptive and relate this to system performance and, in turn, user performance. Like Collopy and Sitterle, Yang et al. note that learning-enabled system performance will fluctuate and diversify with unseen data a system is exposed to over time. As development goals seek increasingly autonomous learning systems with greater adaptive capabilities, the very nature of systems being learning-enabled means the behavior at one point in time may not at all be the behavior at another. In LVC simulations, this poses substantially increased difficulty in accounting for the behavior and time elements.

3.2 What Does it Mean to Team?

When considering how humans and AI-enabled systems can function together to achieve new operational capabilities beyond either alone, one must address what it means to team. Interactions that significantly impact a collective performance dynamic can be intentional or unintentional. Intentional interactions are those considered from the outset in the design and training for elements (human or machine) that interact with other elements to achieve a greater capability. Unintentional interactions are those not pre-considered and explicitly designed for, yet impact performance outcomes none-the-less. For the former, many studies anthropomorphize the entire dynamic and consequently focus on dimensions well-studied for human-human teams: the value of team interactions, trust as a matter of personally ascribed agency, social intelligence, shared understanding and situation awareness, etc. Others eschew "team" and instead focus on "human-AI interaction" design challenges relevant to emergent behaviors; this includes task sharing and design for skill complementarity (Parasuraman et al. 2000; Wilder et al. 2020; Yang et al. 2020). Such approaches typically highlight using the AI-enabled system as a tool, but can be extended to consider a more complex teaming between humans and adaptive, learning-enabled AI systems more analogous to existing human-human dynamics and shared decision making. Multiple systems with heterogeneous capabilities and high autonomy (even if the overall capability is simple and not adaptive) may also interact cooperatively via negotiation or similar algorithmic methods – a current, intentional design paradigm.

For unintentional interactions, producing collective even if not collaborative outcomes remain critical for military scenarios, notably for the increasing use of attritable systems. Elements may not explicitly cooperate, or may function in variable membership groups without direct algorithmic control, but collective actions still create constructive or destructive performance dynamics. This may occur in AI-AI and human-AI interactions and is highly relevant to cyberphysical effects, in which interactions support response to, interaction with, or creation of direct effect(s) in a physical environment (e.g., fires, jamming, etc.).

That being said, this paper is concerned with intentionality and, specifically, how the LVC paradigm may be extended to enable improved outcomes via intentional design and principles discoverable only through the rigorous experimentation offered by the construct beyond all-digital M&S abstractions. Moreover, for intentional interactions, the common views above seem inadequate for the challenge likely to manifest as we strive to advance realizable capabilities in theater. Consequently, in this paper, we define teaming as captured in the introduction: *intentionally developed relationships and forms of interaction specific to humans working with AI-enabled systems designed to work cooperatively with those humans to achieve capabilities beyond either side acting alone*. Teaming embodies intentionality. This definition goes beyond the notion of AI-enabled systems as simple means (i.e., tools or interactive appliances) but does not ascribe human contextualization abilities or emotive characteristics across the hybrid construct. There will be three primary intentional teaming categories for which we will need to develop effective frameworks, foundational studies, and pilot implementations to promote effective design and development of design principles: (i) AI-AI teaming, especially for increasingly learning-enabled, adaptive systems, (ii) human-AI cyberphysical teaming as described above, and (iii) human-AI analytical teaming, where interactions are necessary to add context and promote abductive reasoning about complex problem spaces (e.g., derivation of situational awareness and decision-making support).

3.3 Key Drivers and Functional Groupings Most Relevant to Operational Scenarios

Another aspect critical to understanding human-AI relationship dynamics is the concept of intelligence. Dellerman et al. (2019) and Dubey et al. (2020) each discuss the synthesis of AI and human capabilities, capitalizing the strengths of each to achieve desired results. The groups note that "intelligence" strengths differ across the two, and how they are synthesized is a major determinant of behavioral outcomes. Kaplan and Haenlein (2019) define AI as "*a system's ability to interpret external data correctly, to learn from such data, and to use those learnings to achieve specific goals and tasks through flexible adaptation*." This definition is most in line with the intent of this paper and – importantly – with the concept of teaming defined above. For human-AI teaming in real operational environments, two nuances that should be added are aspects of action, namely scenarios where AI-enabled systems can take action with real world effects, and that AI-enabled systems can operate at speeds and scales across vast amounts of data and dimensions that far exceed human capacity. Humans, for their part, bring powerful abductive reasoning and contextualization. The relationship dynamics are not balanced.

Relatedly, studies emphasize either agency or autonomy with AI systems. Abedin et al. (2022) discusses the significance of tensions between human and AI system agency on outcome dynamics as well as the influence of communication on human-AI collaboration and the ability to achieve a hybrid, or augmented capability. Caldwell et al. (2022) emphasize AI autonomy in human-AI teams and its influence on distributed formulation of decisions across a hybrid team and the role of trust. Often, it seems agency and autonomy are used interchangeably, and so some works diligently differentiate the two (Cummins 2014; Rubel et al. 2021; Bennett et al. 2023). Agency emphasizes capacity to take action, unlike deterministic automation; agency may also take the form of performing tasks in pursuit of objectives specified by a principal (i.e., a human or another AI system), and hence autonomy may be limited. Autonomy is a more comprehensive concept of self-rule, a freedom to act without external control or influence. AI systems may have any of these characteristics. Agency may be within aspects of command and control, whereas autonomy capabilities will more strongly impact the nature of "teaming" with humans.

This relates directly to one of the most frequently cited drivers of effective human-AI teaming – trust. Dimensions of assurance associated with origins, security, asset tracking, authentication, etc. are used to define trust across machine-to-machine and human-machine interactions. However, for humans, trust refers to placing oneself in a vulnerable position, by conferring responsibility to and relying on another entity in the face of uncertainty (i.e., taking a leap of faith). Though related to and often conflated with confidence, confidence derives from belief that events or capabilities occur in line with one's expectations, based on prior experience and clear conveyance of uncertainty or assessment of risk. Confidence can be established through outcomes achieved by conveying trust or by executing control (i.e., over an agent that is not fully autonomous). Adams (2005) articulates the distinctions between the two and their importance related to defense, and the National Academies of Sciences, Engineering, and Medicine (2021) report on human-AI teaming notes that "trust in technology may differ from trust in people" as it expounds on the dimensionality of this concept further. In all cases, trust and confidence are not single-state concepts but dynamically form and break down as events, data, and experiences occur over time.

Many of the aforementioned studies propose frameworks or categorizations whereby future research in human-AI dynamics could be organized and studied. The specifics of each are heavily influenced by how the groups interpret each of the main concepts discussed in this section: teaming, intelligence, autonomy, and trust. In general, however, there are commonalities across key drivers that are germane to the operational problem and map well with the categories defined in the National Academies report: level of automation, AI dynamics and temporality, granularity of control, and other human-AI team interaction issues including trust and bias in both human and AI. Taking an operational lens, we can define a high-level functional distillation that aligns to core warfighting needs to complement the National Academies categorization. These functions are illustrated in Figure 4, although we recognize these processes will often not occur in a linear fashion in real-world application. As AI system and human processes co-evolve, the behavior dynamics in each functional grouping as well as across the entire functional space are highly-non-deterministic. Extending the LVC paradigm may offer a solid foundation through which we can systematically explore, develop, and train for the hybrid capabilities we need operationally.

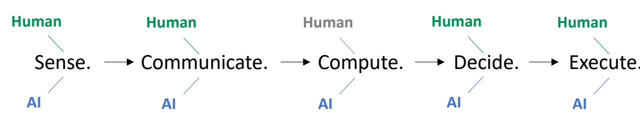

Figure 4: High-level functional distillation of warfighting needs to perform an operation.

4 AI IN SIMULATION

Recent progress with AI and ML in the sciences has resulted in several advancements: import of domain knowledge into ML models and export knowledge back to the scientific domain; leveraging ML for numerically intractable simulation and optimization problems; generating myriads of synthetic data; quantifying and reasoning about uncertainties in models and data; and inferring causal relationships in the data. It is at the intersection of AI and simulation sciences where we can expect significant strides in scientific experimentation and discovery. In Lavin (et al. 2021), they present a unifying holistic perspective to advance the intersection of AI and simulation sciences. They coin this area Simulation Intelligence, and present a roadmap for the development and integration of the algorithms necessary to merge scientific computing, scientific simulation, and AI. This is important work for LVC follow, but it is more narrowly focused on constructive simulations. When considering AI and ML in a broader context, we can think about two general categories to explore: how AI benefits from simulation and how simulation benefits from AI.

4.1 How AI Can Benefit from Simulation

Games and simulations are commonly used as a testbed for developing and training AI/ML algorithms. Games such as StarCraft II, Dota 2, Atari, Go, chess, and heads-up no-limit poker have all been used as platforms for training artificially intelligent agents (Goecks et al. 2022). Fawkes (2017) gives details behind several of these examples. DeepMind tested its AI using computer games, claiming its system was not pre-programmed rather it learned from experience, using only raw pixels of Atari games as data input. In 2013, they published a paper describing an AI playing seven different Atari 2600 video games (Pong, Breakout, Space Invaders, Seaquest, Beamrider, Enduro, and Q*bert). In 2017, OpenAI developed an AI agent that beat professional players in the strategy game Dota 2.

Simulations are also being explored as a medium to train robotic systems. Nvidia created the Isaac simulator in which systems can self-learn through trial and error to carry out tasks and interact with the simulation. Alphabet's Waymo has developed autonomous cars through testing in both the real world and in simulation. In 2016 Waymo logged 3 million miles on real world public streets and 2.5 billion virtual miles in simulation systems. Waymo gave three reasons for using simulation: more miles can be driven than would be possible with a physical fleet; simulated miles focus on difficult interactions for the cars rather than uneventful miles; and the development cycles for the software can be much faster. Lastly, Reinforcement Learning (RL) can also benefit from simulation. Since RL operates on a model of a system, this model can be developed using simulation.

4.2 How Simulation Can Benefit from AI

A recent Dagstuhl seminar, *Computer Science Methods for Effective and Sustainable Simulation Studies*, addressed methodological challenges in conducting effective and sustainable simulation studies. The seminar broke into three working groups, one of which investigated how ML and M&S can be effectively integrated. Initially, two topics were identified: i) AI/ML + Simulation, and ii) Enabling Models to Run Efficiently on Heterogeneous Hardware (Cai et al. 2022). The group eventually merged the discussion into one topic: Intelligent Modeling and Simulation Lifecycle. The discussion focused on how M&S can benefit from advances in AI/ML as well as emerging hardware, computing paradigms and systems.

Current societal and technical challenges require building increasingly complex models and carrying out larger scale simulation experiments, which call for more efficient and intelligent approaches in all aspects of simulation studies, from model creation to model execution and experimentation. The working group examined existing and emerging AI/ML techniques in the context of the M&S lifecycle, and what resulted is shown in Figure 5. Using this lifecycle as a framework, we will discuss a few examples of how AI/ML is being used or envisioned in each phase of the modeling and simulation lifecycle.

4.2.1 Creation

One area of creation is when simulations need agents to reason in more human-like ways, something that may be described as cognitive AI. In this, the decision logic uses factors and reasoning similar to what humans like to believe is the basis of their actual behaviors (Davis and Bracken 2022). The earliest attempt to do this was in 1988 in the SIMNET program with semi-automated forces (SAF) (Shiflett 2013). As larger and more complex exercises were conducted, more vehicles were needed to provide a realistic context for training. However, the original SAF didn't constitute true AI, in that the AI techniques at that time only constitute "automated computation," or a predetermined set of responses to a predetermined set of inputs (Fawkes 2017). Historical examples of AI used in SAF for can be found in (Oswalt and Cooley 2019).

As AI/ML techniques matured, they have been used in numerous simulations to represent behavior. RL is one technique used to provide agents with winning strategies, particularly in combat models. For example, RL has been used in agents to engage in various tactics and to move and attack a defender in a combat simulation (Goecks et al. 2022). As described by Szabo, RL approaches have also been successful at solving problems in dynamic environments and in some cases when dealing with incomplete information (e.g., contested and dynamic environments with poor and unreliable network conditions) (Cai 2022). Another example described by Cai includes using a ML approach to create a car-following model (instead of a traditional physics-based model) and how to dynamically calibrate the model (Cai 2022).

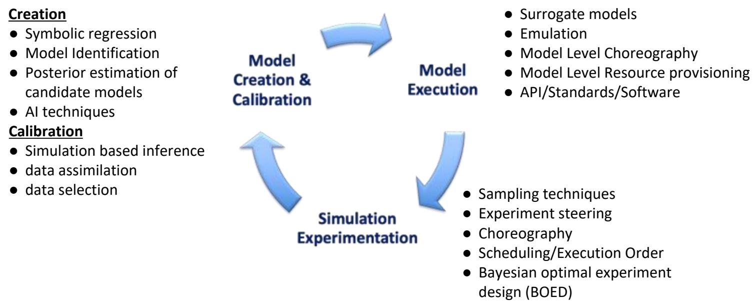

Figure 5: The intelligent modeling and simulation life cycle.

A number of recent papers have postulated how AI/ML capabilities can be incorporated into wargames. One approach modified and augmented the rules and engagement statistics to enable (1) remotely operated and fully autonomous combat vehicles and (2) vehicles with AI/ML-enabled situational awareness, which included their vulnerability to selected enemy countermeasures, such as jamming (Tarraf 2022). While AI decision-making has recently focused on games, modeling decision-making strategies at both tactical and strategic levels requires novel algorithms that can operate within dynamic environments with changing rules, uncertainties, individual biases and randomness (Yuksek et al. 2023).

4.2.2 Calibration

With the emergence of using sensor data as input to simulations and analytics to derive insight from massive data sets, ML techniques can be used to extract useful knowledge and insight from the data to facilitate model development and calibrate simulation models. As discussed by Tan, techniques include simulation-based inference, which link simulation models with empirical data by designing statistical inference procedures, and data assimilation, where the observed data are assimilated into the model to produce a time sequence of estimated system states (Cai 2022).

Several papers discussed the idea of using large volumes of data to provide greater decision speed of ML algorithms for military planning (Goecks et al. 2022). As described in (Yuksek et al. 2023), an intelligent wargaming approach was proposed to evaluate the effectiveness of a military operation plan in terms of operational success and survivability of the assets. The goal was to use AI to discover tactics and suggest effective concepts of use for new military capabilities under consideration. Another paper discussed the application of ML to create algorithms from massive intelligence collection on adversary operations. What once took months or years to collect and analyze may now be available in very short periods of time, as algorithms on operational procedures for wargames or M&S (Davis and Bracken 2022). Oswalt and Cooley (2019) discuss how ML/AI can ingest data and outcomes, and develop (and extend) rules, to reflect a real-time understanding of the battlefield, especially strategy, operations, and tactics. They suggest AI/ML techniques can adapt LVC training systems for things like order of battle and concepts of operation.

4.2.3 Execution

There are many ways AI/ML can support the execution of LVC simulations. Recent research focused on using a data-driven approach to improve performance of simulation execution and simulation-based optimization. The research used ML to dynamically analyze simulation state to determine level of details to be used in the model of an object during simulation execution. The objective was to reduce the simulation runtime while maintaining accuracy of the simulation results (Cai 2022).

AI can also promote better understanding of the delivery, pace, and content of just-in-time training and long-term educational opportunities. Often, training is static; a trainee starts at the beginning of a program and works through all modules on all required topics. An AI enabled simulation-based training system could adapt in real-time to trainees' progress, allowing them to receive more training on deficient tasks and less on highly proficient tasks, thereby reducing time and increasing efficiency (Cooley and Oswalt 2021). In LVC training, AI could ingest data and outcomes, and develop implementing rules that tailor training in numerous simulations for hundreds of participants, across geographic locations (Oswalt and Cooley 2019).

4.2.4 Experimentation

Data analytics and ML techniques can be used to manipulate or steer simulation experiments on the fly. Cai (2022) describes an approach to dynamically predict the usefulness of a simulation run. If the results of a simulation run won't contribute to the overall optimization objective, then the simulation run can be terminated early, thereby reducing the total number of runs required in a simulation-based optimization process. Similarly, ML/AI could optimize networks, reduce latency, and efficiently distribute resource requirements. It could advance the inclusion of environmental parameters in simulation-based training events by providing insights on when, where, and in what context the myriad of possible weather effects significantly impacts force employment outcomes (Oswalt and Cooley 2019).

5 A NEW PARADIGM FOR LVC + AI TEAMING

The military community is traditionally and understandably risk-averse; we must be intelligent about how to design, operate, and train for hybrid human-AI teams. Most studies on human-AI interactions focus on the human dynamic of human-AI teaming. In theater, however, teams will consist of humans and AI systems with heterogeneous capabilities and specialization. Teams will include human-human, human-AI, and AI-AI dynamics with each being bidirectional. Operations require a capability-oriented form of teaming where hybrid teams make decisions and act on them to produce a synergistic if not unified capability with real-world consequences. All levels of sensing, communicating and coordinating, computing, deciding and executing will need to be explored; all may take place in the face of uncertainty and severe time constraints and be distributed to varying degrees depending on the operation.

This is a complex and intractable problem space, one for which LVC may offer the best framework for meaningful and comparative study leading to better, intentional design and well-founded training. There are several possibilities to start accommodating a new AI dimension. After considering whether it was simply part of the Autonomy quadrant, we decided the human-AI teaming interactions are more complex. Thus, we decided it required adding a third dimension to the current planar construct, shown in Figure 6. In other words, we propose to expand LVC to treat AI-enabled systems as another "Actor" in the simulation framework, distinct from real people and simulated platform interfaces (M&S components traditionally employed to add "real feeling" (virtual)) or entirely abstracted versions of likely dynamics (constructive). Fundamentally, we need the LVC construct to capture what it means for a human and AI to team together. Figure 7 is a planar representation, using binary notation to indicate whether a given LVC entity (people, conventional system, or AI system) is real, simulated, or not present.

The concept of "simulated" or "not real" AI is intentionally absent in these figures. Using the existing LVC paradigm as a guide, we considered "simulated AI" or "not real AI" to be AI abstracted and simulated at a lower level of fidelity than the true algorithms (i.e., the "real AI"). While possible to do, it will likely not reflect true performance of an AI system, and the software nature of these elements makes it more reasonable to simply use a copy of the AI software itself rather than a less representative version. Note this is entirely distinct from using simulated data to train AI algorithms. True AI systems (i.e., "real", not some lesser version) may be trained on simulated data and then fielded in the LVC model as-trained or for additional training, which is accounted for through the representation of simulated people and traditional systems together with "real AI" in the construct (i.e., Constructive$_E$ (0, 0, 1) in Figure 7). Accordingly, and as there is no "real" versus "not real" characterization of AI in the literature, these figures reflect the current LVC model as "AI not present" and the proposed extension as "AI present". The extended LVC model explicitly calls out Adaptive, learning-enabled AI systems. As explained in Section 3, narrow AI and ML systems, for which there is no learning or adaptive capability function similar to automated engineering systems, are inherently already accounted for in the conventional LVC model.

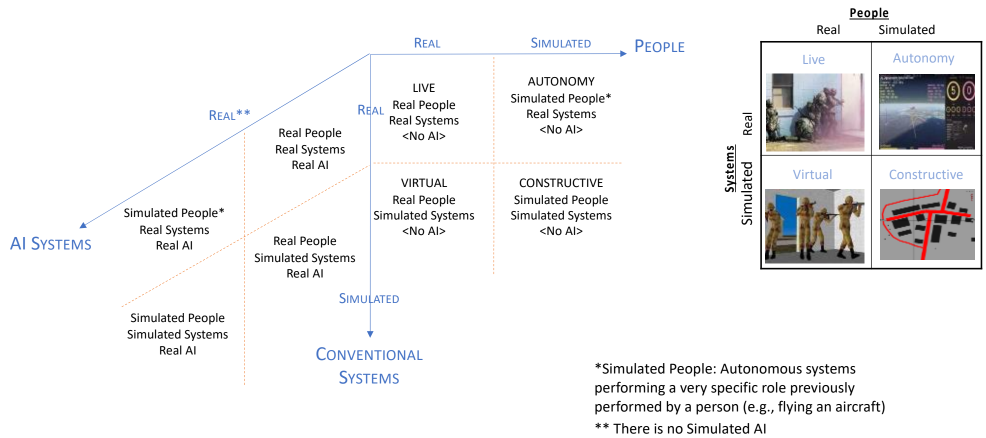

Figure 6: The revised LVC model including artificial intelligence.

	No Adaptive/ Learning-Enabled AI		Real Adaptive/ Learning-Enabled AI	
	Real People	Simulated People	Real People	Simulated People
Real Systems	Live$_C$ (1, 1, 10)	Autonomy$_C$ (0, 1, 10) *(Current blank square)*	Live$_E$ (1, 1, 1)	Autonomy$_E$ (0, 1, 1)
Simulated Systems	Virtual$_C$ (1, 0, 10)	Constructive$_C$ (0, 0, 10)	Virtual$_E$ (1, 0, 1)	Constructive$_E$ (0, 0, 1)
	Conventional LVC		*Extended LVC*	

State:
X (P, S, A)
P – People
S – System (Conventional)
A – Adaptive/ Learning-Enabled AI

Binary representation:
0 – Simulated
1 – Real
10 – Not present

Subscripts:
C – Conventional
E - Extended

Figure 7: Planar representation of revised LVC model.

To illustrate how AI extends our current understanding of LVC, consider several examples of AI defense applications discussed in Frank (2022):

1. Attention Management (Live$_E$)
   Decision-makers face the prospect of information overload. AI can monitor information flows, detect anomalies or events of interest, identify trends, and raise issues for decision-makers.

2. Information Exploitation and Model Validation (Virtual$_E$)
   AI can exploit information to optimize military operations. Successful exploitation depends on a clear sense of the problem to be solved, and efficient and effective routines to address them.

3. Exploratory Analysis (Constructive$_E$)
   Exploratory analysis may provide organizations with deeper insights into adversarial behavior. AI can seek multiple explanations for available intelligence making them less vulnerable to surprise.

4. Autonomy and Principal–Agent Relations (Autonomy$_E$)

> Initial development and employment of autonomous systems and Autonomous Weapons Systems in the battle-space may include robotic vehicles for logistics support and C4ISR systems.

AI-enabled teams will increase the adaptive capacity of military forces. If AI can effectively focus decision-makers on critical problems, perform robust exploration to discover novel innovations, and optimally solve strategic, operational, and tactical conditions, then the tradeoffs and constraints that lie at the heart of strategic decision-making may be radically transformed (Frank 2022). LVC will be key in this evolution. However, experiments need to be well formed and representative of characteristics teams will experience in the real world. This can be accomplished by using AI techniques to improve design and execution of LVC models (as described in the previous section) in experiments to evaluate these new teams.

Similar to the original 90's LVC work, there are tremendous challenges at the seams when trying to bring simulations and AI together in a unified way. Beyond a construct, as provided here, activities such as building scenarios around the expanded dimensions will be critical for stakeholders to understand requirements; frameworks, foundational studies, and pilot implementations will all need to be developed. This will lead to new interoperability challenges and perhaps a commensurate evolution of standards and other approaches to help address them. Organizations like the Simulation Interoperability Standards Organization will be instrumental in this new paradigm and can pioneer design and development principles that will increase the effectiveness of how human and AI-enabled systems can make decisions together.

6 CONCLUSIONS

Ultimately, advances in AI capabilities are achieved through innovative mashups of mathematics and human insight synthesized into algorithmic form. Even so, AI is advancing to the degree that it can be considered a new, emerging technology with tremendous promise and, if used poorly, peril. Newer AI forms are a technology evolution that offer entirely new ways of working together, whether as humans using AI as an "augmentation" or in a human-AI social ecosystem with truly new capabilities. Increasingly adaptive and learning-enabled AI functionality can overcome much of the brittleness inherent in traditional algorithms; importantly, however, it can be just as unpredictable as it is flexible. For confidence in fielding AI alongside and in collaborative relationships with warfighters, we need to design and operate the technology well and safely. This goes beyond simply ascribing computational or policy guardrails, and this is where LVC offers a path to understand how to achieve effective, intentional teaming relationships.

ACKNOWLEDGMENTS

The initial concepts and overall motivation for this work were developed during internal strategic and technical visioning activities supported by the Georgia Tech Research Institute.

REFERENCES

Adams, B. D. 2005. "Trust vs. Confidence". Contract Report CR-2005-203, Human Systems Incorporated: Guelph, Ontario. Defence Research and Development Canada, Toronto. https://apps.dtic.mil/sti/pdfs/ADA630696.pdf

Abedin, B., C. Meske, I. Junglas, F. Rabhi, and H.R. Motahari-Nezhad. 2022. "Designing and Managing Human-AI Interactions". *Information Systems Frontiers 24*(3): 691-697.

Bennett, D., O. Metatla, A. Roudaut, and E. Mekler. 2023. "How Does HCI Understand Human Autonomy and Agency?" *arXiv preprint* arXiv:2301.12490.

Braddock, J.V. and M.R. Thurman. 1993. "Impact of Advanced Distributed Simulation on Readiness, Training and Prototyping". Report of the Defense Science Board Task Force on Simulation, Readiness and Prototyping.

Cai, W., C. Carothers, D. Nicol, and A. Uhrmacher. 2022. "Computer Science Methods for Effective and Sustainable Simulation Studies". Report from Dagstuhl Seminar 22401.

Caldwell, S., P. Sweetser, N. O'Donnell, M.J. Knight, M. Aitchison, T. Gedeon, D. Johnson, M. Brereton, M. Gallagher, and D. Conroy. 2022. "An Agile New Research Framework for Hybrid Human-AI Teaming: Trust, Transparency, and Transferability". *ACM Transactions on Interactive Intelligent Systems (TiiS) 12*(3): 1-36.

Collopy, P., and V. Sitterle. 2019. "Validation of AI-Enabled and Autonomous Learning Systems". Technical Report SERC-2019-TR-017, Stevens Institute of Technology, Systems Engineering Research Center, Hoboken NJ, United States.

Cooley, T. and I. Oswalt. 2021. "Operationalizing Artificial Intelligence in Simulation Based Training". In *Proceedings of the 2021 Interservice/Industry Training, Simulation, and Education Conference (I/ITSEC)* Nov 29th- Dec 3rd, Orlando, Florida.

Cummins, F., 2014. "Agency is Distinct from Autonomy". *AVANT. Pismo Awangardy Filozoficzno-Naukowej* 2: 98-112.

Davis, P.K. and P. Bracken. 2022. "Artificial Intelligence for Wargaming and Modeling". *The Journal of Defense Modeling and Simulation* 15485129211073126.

Dellermann, D., P. Ebel, M. Söllner, and J.M. Leimeister. 2019. "Hybrid Intelligence". *Business & Information Systems Engineering* 61: 637-643.

Dubey, A., K. Abhinav, S. Jain, V. Arora, and A. Puttaveerana. 2020. "HACO: A Framework for Developing Human-AI Teaming". In *13th Innovations in Software Engineering Conference on Formerly known as India Software Engineering Conference.* New York, New York: Association for Computing Machinery.

Fawkes, A.J. 2017. "Developments in Artificial Intelligence: Opportunities and Challenges for Military Modeling and Simulation". In *Proceedings of the 2017 NATO M&S Symposium*, STO-MP-MSG-149: 11-1-11.14.

Frank, A.B. 2022. "Gaming AI without AI". *The Journal of Defense Modeling and Simulation* 15485129221074352.

Goecks, V.G., N. Waytowich, D.E. Asher, S. Jun Park, M. Mittrick, J. Richardson, M. Vindiola, A. Logie, M. Dennison, T. Trout, and P. Narayanan. 2022. "On Games and Simulators as a Platform for Development of Artificial Intelligence for Command and Control". *The Journal of Defense Modeling and Simulation*.15485129221083278.

Gorman, P. 1991. "The Future of Tactical Engagement Simulation". In *The Proceedings of the 1991 Summer Computer Simulation Conference*, 1181-1186. San Diego, California: The Society for Computer Simulation.

Gorman, P.F. 2011. "Learning to Learn: Reminiscences and Anticipation". In *Proceedings, Interservice/Industry Training, Simulation, and Education Conference (I/ITSEC)*, Nov 28th-Dec 3rd, Orlando, Florida.

Interservice/Industry Training, Simulation and Education Conference. 2013. "Fundamentals of Modeling and Simulation Tutorial". *Interservice/Industry Training, Simulation and Education Conference (I/ITSEC)*, Dec 2nd-5th, Orlando FL.

Kaplan, A. and M. Haenlein. 2019. "Siri, Siri, in My Hand: Who's the Fairest in the Land? On the interpretations, Illustrations, and Implications of Artificial Intelligence". *Business Horizons 62*(1): 15-25.

Lavin, A., D. Krakauer, H. Zenil, J. Gottschlich, T. Mattson, J. Brehmer, A. Anandkumar, S. Choudry, K. Rocki, A.G. Baydin, and C. Prunkl. 2021. "Simulation Intelligence: Towards a New Generation of Scientific Methods". *arXiv preprint* arXiv:2112.03235.

National Academies of Sciences, Engineering, and Medicine. 2021. *Human-AI Teaming: State-of-the-Art and Research Needs*. Washington, DC: The National Academies Press.

Oswalt, I., and T. Cooley. 2019. "Simulation Based Training's Incorporation of Machine Learning". In *Proceedings of MODSIM World 2019*, Apr 22nd-24th, Norfolk, Virginia, 45.

Parasuraman, R., T.B. Sheridan, and C.D. Wickens. 2000. "A Model for Types and Levels of Human Interaction with Automation". *IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans 30*(3): 286-297.

Rubel, A., C. Castro, and A. Pham. 2021. "Autonomy, Agency, and Responsibility". In *Algorithms and Autonomy: The Ethics of Automated Decision Systems*, 21-42. New York, New York: Cambridge University Press.

Shiflett, J.E. 2013. "Observations on the Development and Implementation of Distributed Simulation". In *Proceedings of the Interservice/Industry Training, Simulation and Education Conference (I/ITSEC)*, Dec 2nd-5th, Orlando FL, 1F1301.

Tarraf, D.C., J.M. Gilmore, D.S. Barnett, S. Boston, D.R. Frelinger, D. Gonzales, A.C. Hou, and P. Whitehead. 2022. "An Experiment in Tactical Wargaming with Platforms Enabled by Artificial Intelligence". *The Journal of Defense Modeling and Simulation* 15485129221097103.

Wilder, B., E. Horvitz, and E. Kamar. 2020. "Learning to Complement Humans". *arXiv preprint* arXiv:2005.00582.

Yang, Q., A. Steinfeld, C. Rosé, and J. Zimmerman. 2020, April. "Re-Examining Whether, Why, and How Human-AI Interaction is Uniquely Difficult to Design". In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, New York, New York: Association for Computing Machinery.

Yuksek, B., G. Guner, H. Karali, B. Candan, and G. Inalhan. 2023. "Intelligent Wargaming Approach to Increase Course of Action Effectiveness in Military Operations". In *AIAA SCITECH 2023 Forum*, Jan 23rd-Jan 27th, Orlando, Florida, 2531.

AUTHOR BIOGRAPHIES

MARGARET L LOPER is Associate Director for Operations for the Information & Communications Laboratory at the Georgia Tech Research Institute in Atlanta, Georgia. She has been involved in parallel and distributed simulation for more than 30 years, and has taught simulation courses for both academic and professional education. She holds a PhD in Computer Science from the Georgia Institute of Technology. Her email address is margaret.loper@gtri.gatech.edu.

VALERIE B SITTERLE is Chief Scientist for the Systems Engineering Research Division in the Electronic Systems Laboratory at the Georgia Tech Research Institute in Atlanta, Georgia. She has performed design and analysis of systems across operational contexts for over 25 years and serves on the Research Council for the Systems Engineering Research Center. She holds a PhD in Mechanical Engineering from the Georgia Institute of Technology. Her email address is valerie.sitterle@gtri.gatech.edu.