From Hazard Features to Resilient Systems: Bridging Neuroscience and Design in XR-AI for Challenging Environments

Jazmin Collins[1], Poppy McLeod[1], Stephen B. Gilbert[2], Michael C. Dorneich[2], Angelique Taylor[3], Andrea Stevenson Won[1], Derek Spangler[4] and Nina Lauharatanahirun[4]

[1]*Cornell University, Ithaca, NY, 14853, USA*
[2]*Iowa State University, Ames, IA, 50011, USA*
[3]*Cornell Tech, Cornell University, New York, NY, 10044, USA*
[4]*Pennsylvania State University, University Park, PA, 16802, USA*

Abstract

Extended reality (XR) combined with artificial intelligence (AI) offers significant value for teams in challenging environments through three primary applications: on-site tool use (e.g., AR glasses), training simulations, and realistic behavioral research enabled by virtual reality (VR). For all of these needs, realistically representing the salient aspects of hazards is valuable. While existing simulation work has been domain-specific, we propose a systematic approach: a unified framework that bridges neuroscience, ethology, and simulation design to systematically characterize hazards across domains and allow us to create “isometric hazard profiles" of any given threat. By grounding XR-AI systems in how humans actually perceive and respond to threats, this framework enables systems that augment rather than override human expertise in high-stakes contexts, supporting field deployment, team training, and empirical research.

Keywords

virtual reality, hazards, teamwork, simulation, fear, behavior

1. Introduction

Extended reality (XR) has value for teams in challenging environments in at least three ways. First, it can be a tool used on-site through applications such as augmented reality (AR) glasses. Second, it can be used to create training simulations. Finally, it can be used to study teams through the ability of virtual reality (VR) to elicit realistic behavior.

For all of these needs, realistically representing the salient aspects of hazards is valuable. However, current hazard frameworks are critically fragmented across domains, with many disciplines identifying their own terms and definitions for hazards that can make it difficult for VR developers to reliably implement or communicate about them. Thus, we argue for a universal behavioral language for hazards—grounded in neuroscience—to enable the design of resilient XR systems that align with human evolutionary threat perception.

Beyond improving XR experiences, such a language could improve how AI systems integrated with XR tools offer guidance and interventions, better helping human-AI teams proactively recognize, react, and manage hazard response in chaotic, dangerous environments. However, to realize future human-AI collaboration for hazard management, we must first identify the hazard features that are commonly encountered in high-stress contexts. Such features can then be used by AI technologies in XR environments to model and anticipate human responses (e.g., errors, stress, reaction time, etc.) to environmental threats, optimally informing interventions which improve the speed and accuracy of hazard responses.

XR for Challenging Environments - Enabling Human Performance and Agency under Stress – 1st Workshop edition at CHI Conference on Human Factors in Computing Systems (XR4CE 2026), April 13-17, 2026, Barcelona, Spain

jc2884@cornell.edu (J. Collins); plm29@cornell.edu (P. McLeod); gilbert@iastate.edu (S. B. Gilbert); dorneich@iastate.edu (M. C. Dorneich); amt298@cornell.edu (A. Taylor); asw248@cornell.edu (A. S. Won); dpspang@psu.edu (D. Spangler); nina.lauhara@psu.edu (N. Lauharatanahirun)

To this end, we propose a conceptual framework for parameterizing hazards by synthesizing multidisciplinary perspectives on hazards. We discuss related work from hazard characterization in risk communication, neurobiological responses underlying threat perception, and hazard parameterization in video-game/VR/AR design. We describe initial steps for implementing a hazard system in a simulation context via **isometric hazard profiles**: collections of universal features of hazards that characterize equivalently-threatening hazards across domains and risk contexts. Finally, we indicate potential challenges that could be fruitfully addressed in this workshop.

2. Related Work

Previous researchers and organizations have examined ways to characterize hazards in a variety of scenarios [1, 2, 3, 4, 5]. For example, the UN Office for Disaster Risk Reduction classified hazards via the physical processes causing them (e.g., a storm caused by meteorological phenomenon is different from a disease brought on by biological processes), as well as certain aspects defining the level of danger a hazard poses, such as its risk of cascading into further issues [2]. This aligns with work by the Centre for Research on the Epidemiology of Disasters [4] and the Intergovernmental Panel on Climate Change [3]. Other researchers have similarly attempted to define hazards across multiple domains by examining triggering mechanisms (i.e., what is causing hazards) as well as which sequential behaviors hazards move through as they build into larger threats [5].

These works provide overarching typologies through which we can identify hazards via their human-facing features. Thus, we do not focus on larger classifications of hazards (e.g., geophysical, meteorological, biological, technological), but on granular characteristics of the hazards' behaviors, such as intensity, duration, and triggering mechanisms. By drawing out these particular mechanisms which describe how hazards act, we aim to develop a framework that can use such terms to wholly describe any hazard's behavior.

We also propose that hazard features can be partly categorized based on their biological and evolutionary significance, thus drawing from neuroscience, ethology, and ecological sciences [6, 7, 8, 9, 10, 11]. Biologically speaking, organisms (including human actors) are embedded in complex ecologies filled with danger and uncertainty. These dangers are distributed across multiple spatiotemporal scales, and it is imperative for the actor to anticipate, appraise, react to, and remember where/when those risks occurred—otherwise, survival is at risk.

Much work in the areas of ethology and ecology has used an anti-predator model that informs similar anti-threat responses in humans [6]. They have documented a “landscape of fear" in which the animal considers the properties of predator stimuli and their relationships to the surrounding environment using an iterative “risk assessment" process [7, 8, 9]. This investigative process allows the animal to mount a strategic defense (fight, flight, freeze, and more) that optimizes survival—in other words, a specific set of neural, physiological, and behavioral responses tuned to the situation. Those threat responses often form so-called “defensive states" that might be called fear, anxiety, and aggression in non-human and human animals [12, 13]. These anti-predator responses, and the risk assessment that determines them, are often reflexively activated, primordial, and evolutionarily rooted across species [6, 12, 14]. Anti-predator defense therefore represents a translational foundation for other defensive responses to different threat types, not just predators. This includes weapons, other human beings, and more symbolic threats like challenges to self-esteem [10, 11].

By exploring literature in the domain of anti-predator responses and fear landscapes, we aim to identify hazard features that represent ambiguous sources of threats, such as social threats from other humans or internal threats from a user's own mental model. By identifying the language used to describe these neuroscience-driven reactions to threats, for instance, we can anticipate and model elements of a user's response to hazards within our framework, such as a user's cognitive load, emotional states, or stress levels when facing more agentic hazards.

Finally, to guide our implementation of realistic hazards, we look to XR modeling and video game design. When designing video games and XR environments, designers must decide where to emphasize fidelity (e.g., whether to focus on photorealistic visuals or on believable scenarios). The Interaction Fidelity Model [15] delineates eight fidelity types, with the most important for hazard parameterization being simulation fidelity (does the environment react appropriately to users' actions) and experiential fidelity (does the user feel the same way about the environment as they would in the real-world equivalent). Experiential fidelity is additionally linked to XR coherence [16] and authenticity [17], which measure how well the environment aligns with user expectations. In game design, researchers have used procedural content generation [18] to generate games or game levels of a specific difficulty [19, 20], as well as evolutionary algorithms like MAP-Elites to generate game enemies that players perceive as easy, medium, and hard [21]. The challenge with algorithmic approaches like these is defining the best behavioral descriptor factors from the start; the results depend heavily on the choices of the designer. When designing hazards, we propose these descriptor factors can be best chosen from the interdisciplinary space of neural and psychological perception of hazards.

3. Current Work

Building on the literature reviewed above, we have begun to investigate the specific features of hazards that give rise to different threat response strategies such as decisions to "fight," "flee," or "freeze." Such features may be fruitful starting points for categorizing hazard features in high stress operational environments. First, the **physical** features of hazards—such as size, shape and distance—contribute to the perception of threats and thus, human threat responses (cognitions and behaviors). Second, the **spatiotemporal** features (when and where hazards occur) influence how humans interpret these physical features. Finally, how features are **perceived** also influences human responses. For example, the level of ambiguity (or perceptual identifiability) of an environmental hazard and the intensity of the stimulus both can influence response. Importantly, these areas are not exhaustive of all categories of hazard features–they represent an initial starting point that we can use to implement a functional version of this framework (detailed below). Future additions to our framework, particularly those represented by non-agentic hazards, are discussed in Section 4.

Below, we describe an initial effort to parameterize a set of hazard "behaviors." Collectively, these behaviors form a descriptive **isometric hazard profile** that tells us how a hazard should act. The following profile parameters have been implemented in a Universal Hazard Controller prototype developed in Unity. This Controller allows developers to drag-and-drop hazards from their Unity environments into slots in the Controller script, and use simple interfaces such as sliders and numeric entry fields to set the value of the parameters. Based on the values selected for a particular hazard, its behavior in the environment will be altered. We provide examples of how this system works with two very different sets of hazards: marine animal hazards and structural fires.

- **Aggression**: A parameter that defines how likely a hazard is to trigger more aggressive behaviors (e.g., how likely a shark is to attack the player rather than flee, how likely a fire is to spread quickly rather than die out).
- **Caution**: A parameter that defines how likely a hazard is to linger on certain behaviors before moving to others (e.g., how likely a shark will circle something before swimming closer to investigate, how likely a fire will stay burning on one object before spreading to another).
- **Patrolling**: A behavior where a hazard has a given territory that it keeps to or starts within, and simply moves around that territory (e.g., a shark circling through a reef, a fire spreading within a single room).
- **Approaching**: A behavior where a hazard moves or expands towards a player or target in a scene (e.g., a shark swimming closer to something that caught its attention, a fire spreading when a door or window is open).
- **Attacking**: A behavior where a hazard actively collides with a player or target in a scene (e.g., a shark lunging at a player's body and "biting" it, a fire touching and lighting up the object beside it).

Ultimately, these isometric hazard profiles allow us to compare the behaviors that hazards display across domains, noting how similar and dissimilar they are under various conditions. Additionally, we can identify hazards that pose equivalent threats across domains; for example, a Patrolling shark and a Patrolling fire pose the equivalent threat of a hazard that may expand or move if given the chance, but is currently contained to one location.

4. Opportunities, Challenges and Next Steps

By grounding hazards in a universal lexicon, we enable resilience by design. This shared language enables VR developers to think critically about how they introduce hazards in challenging simulations from the beginning, considering various features of hazards that can be connected to risk literature. This makes the process of generating hazardous environments more resilient and grounded from the beginning stages. Further, through the isometric hazard profiles that can be generated by this shared lexicon, developers can create hazards that are explainable and transferable between disciplines, making simulation development more resilient for collaboration and expansion.

There are also more technical benefits of this framework to simulation development. A system that can model threat features matching human neurobiology can improve the use of XR in challenging environments in multiple ways. In the context of human-AI teaming, AI interventions that respect evolutionary risk assessment processes may feel intuitively correct rather than alien or disruptive. For example, an AI system that provides distal alerts for ambiguous threats (allowing time for human risk assessment) versus immediate directives for proximal, high-intensity threats aligns with the freeze-versus-flee response patterns hardwired into human physiology. This biological alignment means operators don't need to consciously override their instincts to follow AI guidance; instead, the AI augments existing threat perception circuits.

In addition, our implementation of universal hazard characteristics through simple, public-facing interfaces like drag-and-drop or sliders lessen a substantial workload that simulation designers face when creating hazardous simulations. This work would allow for rapid generation of realistic, customizable hazards with minimal effort in the future. For instance, we propose a potential AI-powered hazard profiler, Hazardrama, that could automatically generate recommended isometric hazard profiles for the intended hazards in a developer's environment. If a developer is working on a simulation of a forest with various wild animals, they could input the base physical variables that their animals are meant to have (e.g., speed, size). The profiler can than suggest different optimal settings for that hazard; if a wolf in the forest is injured but hungry, then it will be recommended a High Aggression / High Caution profile, whereas if it is healthy and defending a territory, it might have High Aggression / Low Caution. This proposed application would allow developers to procedurally generate new and variable hazards for any type of simulation.

Such procedurally-generated hazards also prevent overfitting to specific scenarios. Training that systematically varies ambiguity (lighting conditions), predictability (timing of threat onset), and behavioral patterns (patrolling territories) builds cognitive flexibility rather than rote memorization. Teams learn to recognize threat signatures—combinations of features that indicate danger—rather than specific instantiations. This prepares them for the fundamental unpredictability of real challenging environments, where no two emergencies are identical but underlying threat dynamics follow recognizable patterns. The lexicon also enables cross-domain communication: firefighters discussing "high ambiguity, low predictability" scenarios can immediately connect with combat medics facing similar isometric hazard profiles in different contexts.

However, our work still faces many limitations. At the moment, our implementation of a universal VR hazard simulation is still in development, and lacks representation of all areas of hazard research, particularly non-agentic work. Hazard interpretation relies heavily on contextual information, as hazards may lack visible or physical attributes and rarely occur in isolation. Instead, they emerge within complex, dynamic environments that span both social and non-social contexts. Threat responses can be self-directed or other-directed, with the latter potentially targeting individuals at varying social distances from the actor. Hazard perception is also known to be significantly affected by experience with the hazard, as documented in driving [22], machining [23], and disaster preparation [24], among others. Also, multiple previous studies have explored the impact of personality factors (such as the Big Five factor neuroticism) on risk perception [25, 26, 27]. These results suggest that parameters such as domain expertise and personality factors must also be included in a framework for parameterized hazard generation. To generate a truly complete framework for hazards, we must establish features that account for additional families of hazards, including:

- **Contextual Hazards**: “Invisible" threats such as risks of contamination or widespread equipment failures that have less agentic forms. We may incorporate this through a “Spreading" behavior to represent acting maliciously on multiple inter-connected actors or objects.
- **Self-Generated Hazards**: Internal threats introduced by a user’s own personal and mental profile, such as their fatigue, stress levels, or personality-driven response to threats. We may incorporate this through a “Cognitive Load" behavior that enhances the likelihood of multiple hazards being present at a single time.

5. Conclusion

We propose a robust framework for universal hazard design that establishes a series of parameters (i.e., isometric hazard profiles) which can effectively represent the salient behavior of hazards in the context of teamwork. By developing this framework, we create a common language which simulation designers, developers, teams training for real-life hazards, and other members of the hazard and risk community can utilize to fully explain any hazard. In addition, we propose a potential implementation of this framework whereby a simulation designer can customize realistic hazards by altering a set of hazard behaviors drawn from various fields such as biology, neuroscience, and risk management, using simple interfaces such as sliders. This implementation supersedes the need for designers to create complex code for new hazards, and allows them to create any desired hazard by adjusting universal parameters behind hazardous behaviors. This work not only links together core work on understanding hazards and simulating them across disciplines, but opens the path to developing more realistic hazardous simulations in XR for some of XR’s most critical use cases.

References

[1] I. Burton, The environment as hazard, Guilford press, 1993.
[2] U. N. I. S. for Disaster Reduction (UNISDR), Global assessment report on disaster risk reduction (gar), 2019.
[3] I. P. on Climate Change (IPCC), Special report on managing the risks of extreme events (srex), ???? URL: https://www.ipcc.ch/report/srex/.
[4] C. for Research on the Epidemiology of Disasters (CRED), Classification glossary: Definitions of disaster types, ???? URL: https://doc.emdat.be/docs/data-structure-and-content/glossary/.
[5] J. C. Gill, B. D. Malamud, Reviewing and visualizing the interactions of natural hazards, Reviews of geophysics 52 (2014) 680–722.
[6] M. Kavaliers, E. Choleris, Antipredator responses and defensive behavior: ecological and ethological approaches for the neurosciences, Neuroscience & Biobehavioral Reviews 25 (2001) 577–586.
[7] M. S. Palmer, K. M. Gaynor, J. A. Becker, J. O. Abraham, M. A. Mumma, R. M. Pringle, Dynamic landscapes of fear: understanding spatiotemporal risk, Trends in Ecology & Evolution 37 (2022) 911–925.
[8] S. L. Lima, L. M. Dill, Behavioral decisions made under the risk of predation: a review and prospectus, Canadian journal of zoology 68 (1990) 619–640.
[9] D. C. Blanchard, G. Griebel, R. Pobbe, R. J. Blanchard, Risk assessment as an evolved threat detection and analysis process, Neuroscience & Biobehavioral Reviews 35 (2011) 991–998.

[10] P. J. Lang, M. M. Bradley, B. N. Cuthbert, Emotion, motivation, and anxiety: Brain mechanisms and psychophysiology, Biological psychiatry 44 (1998) 1248–1263.

[11] D. C. Blanchard, A. L. Hynd, K. A. Minke, T. Minemoto, R. J. Blanchard, Human defensive behaviors to threat scenarios show parallels to fear-and anxiety-related defense patterns of non-human mammals, Neuroscience & Biobehavioral Reviews 25 (2001) 761–770.

[12] P. J. Lang, M. Davis, A. Öhman, Fear and anxiety: animal models and human cognitive psychophysiology, Journal of affective disorders 61 (2000) 137–159.

[13] D. C. Blanchard, Translating dynamic defense patterns from rodents to people, Neuroscience & Biobehavioral Reviews 76 (2017) 22–28.

[14] Y.-T. Tseng, B. Schaefke, P. Wei, L. Wang, Defensive responses: behaviour, the brain and the body, Nature Reviews Neuroscience 24 (2023) 655–671.

[15] M. Bonfert, T. Muender, R. P. McMahan, F. Steinicke, D. Bowman, R. Malaka, T. Döring, The interaction fidelity model: A taxonomy to communicate the different aspects of fidelity in virtual reality, International Journal of Human–Computer Interaction 41 (2025) 7593–7625. doi:10.1080/10447318.2024.2400377.

[16] R. Skarbez, F. P. Brooks, M. C. Whitton, Immersion and coherence: Research agenda and early results, IEEE Transactions on Visualization and Computer Graphics 27 (2021) 3839–3850. doi:10.1109/TVCG.2020.2983701.

[17] S. Gilbert, Perceived realism of virtual environments depends on authenticity, Presence 25 (2016) 322–324. doi:10.1162/PRES_a_00276.

[18] N. Shaker, J. Togelius, M. J. Nelson, Procedural content generation in games, Springer, 2016.

[19] L. Climent, A. Longhi, A. Arbelaez, M. Mancini, A framework for designing reinforcement learning agents with dynamic difficulty adjustment in single-player action video games, Entertainment Computing 50 (2024) 100686. URL: https://www.sciencedirect.com/science/article/pii/S1875952124000545. doi:https://doi.org/10.1016/j.entcom.2024.100686.

[20] A. Baldwin, S. Dahlskog, J. M. Font, J. Holmberg, Mixed-initiative procedural generation of dungeons using game design patterns, in: 2017 IEEE Conference on Computational Intelligence and Games (CIG), 2017, pp. 25–32. doi:10.1109/CIG.2017.8080411.

[21] B. M. F. Viana, L. T. Pereira, C. F. M. Toledo, Illuminating the space of enemies through map-elites, in: 2022 IEEE Conference on Games (CoG), 2022, pp. 17–24. doi:10.1109/CoG51982.2022.9893621.

[22] G. Underwood, Visual attention and the transition from novice to advanced driver, Ergonomics 50 (2007) 1235–1249. URL: https://doi.org/10.1080/00140130701318707. doi:10.1080/00140130701318707.

[23] V. G. Duffy, Effects of training and experience on perception of hazard and risk, Ergonomics 46 (2003) 114–25. doi:10.1080/00140130303524.

[24] J. P. Nicholas, A. Donovan, C. Oppenheimer, Experts at risk: The influence of expertise on conceptualising multi-hazard risk perception and preparedness in squamish, canada, International Journal of Disaster Risk Reduction 117 (2025) 105208. URL: https://www.sciencedirect.com/science/article/pii/S2212420925000329. doi:https://doi.org/10.1016/j.ijdrr.2025.105208.

[25] M. Bouyer, S. Bagdassarian, S. Chaabanne, E. Mullet, Personality correlates of risk perception, Risk Analysis 21 (2001) 457–466. URL: https://onlinelibrary.wiley.com/doi/abs/10.1111/0272-4332.213125. doi:https://doi.org/10.1111/0272-4332.213125.

[26] B. Chauvin, D. Hermand, E. Mullet, Risk perception and personality facets, Risk Analysis 27 (2007) 171–185. URL: https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1539-6924.2006.00867.x. doi:https://doi.org/10.1111/j.1539-6924.2006.00867.x.

[27] A. Fyhri, A. Backer-Grøndahl, Personality and risk perception in transport, Accident Analysis Prevention 49 (2012) 470–475. URL: https://www.sciencedirect.com/science/article/pii/S0001457512001066. doi:https://doi.org/10.1016/j.aap.2012.03.017, pTW + Cognitive impairment and Driving Safety.