Advancing Public Literacy of Uncertainty in Science (APLUS) in the Context of Simulation-based North Atlantic Storm Forecasting

Project Rationale

Severe weather events represent important opportunities for public science engagement. North Atlantic storms disrupt lives and incur substantial financial burdens on communities. Hurricanes, tropical cyclones powered by warm ocean waters, and Nor'easters, extratropical cyclones forming as cold Arctic air meets warm moist air from the Atlantic Ocean, dominate the weather cycle from June through April. Over 20 storms of each type can form annually, with several causing significant damage. The 2024 hurricane season, marked by five major hurricanes striking the U.S. mainland, is estimated to have caused $500 billion in damage [1], with impacts extending as far inland as mountainous North Carolina, a location previously advertised as one of the safest areas from climate-driven natural disasters [2]. With storm intensity and frequency increasing due to a warming climate [3], promoting forecasting literacy has become an urgent priority for communities along the East Coast, Gulf Coast, and Puerto Rico [4].

The public can now access more real-time weather data and forecast information than ever before. With real-time weather updates now accessible on smartphones, individuals can monitor minute-by-minute changes during extreme events [5]. As a result, public engagement with weather information is at an all-time high, and weather and forecast communication presents a prime opportunity for increasing the public's engagement with scientific data and science practices through scenarios relevant to their lives. Getting this communication right is crucial. Without a clear understanding of the uncertainties inherent in storm forecasting, individuals can misinterpret forecasts, leading to misunderstandings, poor decision-making, and even mistrust in the science underlying weather forecasting overall [6].

Yet fostering understanding of forecast uncertainties poses persistent and problematic challenges. The proliferation of weather information presents a double-edged sword for public communication. On the one hand, members of the public have increasingly wide access to nuanced information about forecasts, including information about weather models and intricate visualizations such as cones of uncertainty and spaghetti plots [7]. On the other hand, they receive limited if any guidance on how to interpret the probabilistic nature of such visualizations [8]. Instead, individuals end up responsible for perceiving risks and making preparedness decisions based on their own interpretations of these visualizations [9], which can vary widely [10]. This inconsistency poses a significant challenge for weather authorities tasked with effectively communicating storm risks [11], [12], raising public awareness [13], and supporting decision-making during extreme weather events [14].

Fortunately, simulations show promise for supporting public understanding of uncertainty involved in forecasting. Helping the public gain insight into the storm forecasting process, where uncertainties are expected and critical, is essential [15]. Simulations can increase understanding of complex Earth systems, including those that produce natural hazards [16], by illustrating how parameters combine to produce outcomes and how randomness can lead to large-scale phenomena [17], [18]. Recent work by project staff highlights the potential of simulations to enhance understanding of causes and risks in natural hazards [19], [20], with results showing improved comprehension of forecast uncertainty visualizations for storm-based weather events.

This project will design and test a simulation-based experience to enhance public literacy related to storm forecasting. The simulation experience will replicate key aspects of the storm modeling and forecasting practices used by scientists. Through interactive exploration of environmental factors influencing storm paths and the challenges of setting initial conditions and predicting future trajectories, participants will engage with inherent uncertainties in North Atlantic storm forecasting, including modeling, ensemble, and temporal uncertainties. This simulation experience is hypothesized to improve participants' interpretation of probabilistic storm visualizations and storm risk perception.

Project Goals and Objectives

The Concord Consortium (CC), in partnership with the American Meteorological Society (AMS) and Physics Front (PF), proposes the Advancing Public Literacy of Uncertainty in Science (APLUS) project in the context of North Atlantic storm forecasting. The goals of this three-year project are to design a web-based interactive experience and to evaluate its impact on adult participants' Uncertainty Literacy in Atlantic Forecasting (ULAF). These goals will be achieved through the following objectives.

Objective 1: Develop instruments to assess three constructs that constitute ULAF: (1) interpreting probabilistic visualizations of forecasted tracks for tropical storms and Nor'easters (e.g., cone of uncertainty graphics and spaghetti plots), (2) understanding the uncertainties inherent in these visualizations (i.e., modeling, ensemble, and temporal uncertainties), and (3) perceiving risks conveyed by the visualizations. By surveying 400 adult participants, we will validate these instruments and establish a general landscape of public forecasting uncertainty literacy.

Objective 2: Conduct design-based research to develop and refine the North Atlantic Storm (NAS) Explorer, a simulation for modeling storm tracks, with 10 focus group participants in Year 1 and the simulation-based forecasting experience with 100 adult participants in Year 2.

Objective 3: Evaluate the impact of the simulation forecasting experience on participants' ULAF via a pilot study using a randomized-control-trial (RCT) approach. We will recruit a total of 300 participants from communities vulnerable to the impacts of hurricanes and Nor'easters through AMS social media and community engagement programs, including assistance from broadcast meteorologists and weather influencers. We will randomly assign participants to one of two conditions. A control group will examine weather authorities' visualized forecasting information and communicate their ideas about storm risks in message boards, a process designed to replicate the current public communication pathway of storm information on the Internet and social media. A treatment group will engage in an augmented version of this cycle including a simulation forecasting experience with the NAS Explorer.

Objective 4: Disseminate the NAS Explorer to informal educators, researchers, scientists, weather authorities, and the public. We will also distribute research instruments, design frameworks, and communication strategies promoting public ULAF at conferences and through journal publications.

As a Type 4: Integrating Research and Practice project, APLUS directly addresses the stated and required goals of the Advancing Informal STEM Learning (AISL) program as follows:

AISL Goal #1: Learning STEM in informal experiences and environments. APLUS will engage the general public in a simulation-based modeling process of North Atlantic storms as they seek forecasting information online. Participants interactively manipulate factors influencing storm tracks and observe the effects visualized as the simulated storm progresses. By creating multiple storm tracks, they gain insights into the uncertainties inherent in real-world storm forecasting. This experience occurs outside traditional educational settings, helping participants connect their understanding of uncertainty to real-world challenges, such as interpreting storm visualizations provided by weather authorities.

AISL Goal #2: Advancing the knowledge base of informal STEM learning. APLUS will advance the knowledge base by developing instruments to measure uncertainty literacy in Atlantic storm forecasting, designing a simulation-based forecasting experience tailored for risk communication, and evaluating its relevance and impact. See the logic model (Table 1).

Table 1. Project logic model

Objectives	Inputs	Activities	Short-term Outcomes	Long-term Outcomes
1. Instrument development	Literature; expert feedback; survey design tools	Design, test, and refine assessment instruments	Validated instruments measuring ULAF	Standardized tool for assessing forecasting literacy
2. Design-based research	Focus groups; prototype simulations; forecasting experience; participant feedback	Iteratively design, test, and refine NAS Explorer and simulation experience	Improved functionality and usability of the simulation; simulation forecasting experience	Scalable and transferable tools, design frameworks, and strategies for broader applications
3. RCT pilot study	Participant recruitment; randomized assignment	Test the causal impact of the simulation experience	Evidence of improving ULAF	Empirical support for simulation-based informal interventions
4. Dissemin-ation	Conference presentations; academic and trade publications; open-access repositories/ licenses	Share findings and distribute the simulation, experience design, and data	Increased awareness of informal educators; increased tools/data access for researchers and the public	Sustained integration of tools and findings in corpus of literature on using simulations to teach informally

AISL Goal #3: Broadening participation in STEM. APLUS will engage vulnerable communities along the U.S. Atlantic and Gulf Coasts, areas frequently impacted by hurricanes and Nor'easters. Recruitment for instrument development, design-based research, and evaluation will emphasize inclusivity, leveraging partnerships with community organizations and AMS outreach channels. Strategies will ensure the NAS Explorer and simulation experience are accessible to participants regardless of gender, race, socioeconomic status, or education level. Feedback mechanisms will integrate participant input into the iterative design process, and instruments will prioritize accessibility. Advisory Board members with extensive experience in outreach and informal education will monitor the project, ensuring alignment with its goals and enhancing the experience based on participant feedback.

AISL Goal #4: Intentionally community/participant driven. APLUS will address the critical need for better communication of scientific uncertainty related to storm forecasting information in the coastal areas vulnerable to storm impacts. Recognizing that storm forecasting uncertainty is difficult to effectively communicate through text alone, APLUS leverages interactive simulations that integrate user input with dynamic, visualized outputs. This approach empowers participants to actively engage with the forecasting process by allowing them to make changes to environmental factors and observe the resulting outcomes, fostering a sense of agency [21]. APLUS thus moves beyond passive information delivery, enabling deeper comprehension of storm forecasting uncertainty and its implications.

AISL Goal #6: Support learners' participation in and understanding of STEM practices. Through a simulation-based forecasting experience, APLUS will support participants' active engagement with the streamlined version of the meteorologists' computational modeling practice to forecast future storm tracks to grasp uncertainties inherent in storm visualizations and risk communication.

Results from Prior NSF Support

Since 2012, **PI Lee** and CC **Co-PI Pallant** have carried out seven NSF-funded projects to develop curriculum, teaching, and assessment resources including 14 interactive computational models (i.e., simulations) of various environmental, geological, and natural hazard systems. **Co-PI Pallant** led the development of these simulations and associated activities for middle and high school students [19], [22], [23], [24]. **PI Lee** is a recognized leader in incorporating scientific uncertainty into simulation-based modeling and argumentation practices in science education, having developed innovative design, assessment, and research frameworks [25], [26], [27], [28]. **PI Lee** and PF **Co-PI Gweon** have collaborated to create analytic frameworks and advanced log analysis algorithms to study users' interactions with simulations as they engage in scientific inquiry to resolve uncertainties [28], [29], [30], [31]. AMS **Co-PI Price** has led several NSF-funded AISL projects to create and research informal learning opportunities for participant-driven or community-oriented science learning [32] using visualizations [33], [34]. Below, we highlight five key projects relevant to APLUS.

GeoHazard: Modeling Natural Hazards and Assessing Risks (PI: **Pallant**; Co-PIs: **Lee**, McAuliffe, McDonald, Larson; DRL-1812362; $2.8M; 2019-2023). ***Summary of results:*** The project developed curriculum modules for wildfires, floods, and hurricanes, accompanied by simulations and assessment instruments to measure students' understanding of natural hazards and their ability to construct risk-based scientific arguments involving uncertainty. The reliability of the assessment instruments was demonstrated with Cronbach's alpha values of 0.86, 0.87, and 0.91 for wildfires, floods, and hurricanes, respectively. Significant pre-post learning gains were observed across all modules, with effect sizes of 0.69 SD for hurricanes, 0.54 SD for floods, and 0.77 SD for wildfires. Research found simulation interactions accounted for the largest impact on students' pre to post changes in their understanding of natural hazards, risks, and mitigations. In the context of hurricanes, students engaged with uncertainty through various representations, including the cone of uncertainty, simulations, and data graphics. A paired samples t-test revealed a significant improvement in students' understanding of cone of uncertainty visualizations from pre-test to post-test, with an effect size (ES) of 0.31, $t(198) = 4.38$, $p < 0.001$. ***Intellectual merit:*** The project provided design knowledge on developing and integrating simulations into natural hazard curricula; developed new techniques for analyzing student interactions with simulations; and created a pedagogical framework for teaching natural hazards. ***Broader impacts:*** Since public release, the wildfire module has been implemented by 232 teachers and 11,411 students, the hurricane module by 208 teachers and 9,484 students, and the flood module by 127 teachers and 4,593 students. ***Publications:*** 4 published papers, 1 submitted paper, and 5 newsletter articles. See References.

InquirySpace 2 (IS2): Broadening Access to Integrated Science Practices (PI: Dorsey; Co-PIs: Damelin, **Lee**, **Gweon**, Tinker; DRL-1621301; $4.5M; 2016-2022). ***Summary of results***: The project developed three high school modules for data-intensive, inquiry-based, independent scientific investigations in physics, chemistry, and biology. ***Intellectual merit:*** The project developed a pedagogical model of experimental inquiry, a new theory on epistemic engagement as uncertainty resolution during scientific experimentation with physical setups and simulations, and an instrument to measure epistemic knowledge development. ***Broader impacts:*** 50 teachers and 3,000 students used IS2 modules. ***Publications***: 7 published papers and one submitted paper. See References.

STTR Phase I: Cloud-Based Pluggable Learning Analytics Engine for Educational Games (PI: **Gweon**; 1549811; $225K; 2016-2018). ***Summary of results:*** Two learning analytics algorithms were refined and developed to analyze users' simulation interactions: Monte Carlo Bayesian Knowledge Tracing (MC-BKT) [30] and Dynamic Entropy Tracing [35]. ***Intellectual merit***: MC-BKT algorithm works with both binary and continuous scoring of simulation-based educational games, produces distinctive knowledge growth patterns, works in real time as formative assessment in real-world classrooms. ***Broader impacts:*** 364 students and 5 teachers in 3 high schools participated in the project. MC-BKT plugins were released on GitHub. ***Publications:*** 1 online user manual. See References.

AISL: Two Eyes, 3D: Studying Stereoscopic Representations in Informal Learning Environments (PI: **Price**; DRL-1114645; $660,487; 2011-2015). ***Summary of results:*** This was an experimental study of scientific visualizations at a planetarium. ***Intellectual merit:*** We found stereoscopy did not impact immediate learning but did have an impact on delayed recall, likely through increased emotional engagement of the visitor. ***Broader impacts:*** 498 adult visitors and 261 children were surveyed across two studies. ***Publications***: Two published papers. See References.

AISL: Investigating How Museum Experiences Inform AISL-funded Youths' STEM Career Awareness and Interest (PI: **Price** [2019-2022]; due to employer change, current PI is L. Applebaum [2022-2025]; DRL-1906954; $1.2M; 2019-2025). ***Summary of results:*** An experimental study measured the impact of experience with a human patient simulator on awareness of health careers and community health issues on students. ***Intellectual merit:*** Human patient simulators are used in modern medical education programs. This study examined if such simulations could provide a more authentic healthcare experience to secondary school students. ***Broader impacts:*** 1,299 students from 34 schools participated in a mixed-methods, sequential delayed research design. ***Publications*****:** A publication is in preparation. Multiple poster updates have been presented at informal science conferences (Visitor Studies Association, Association of Science and Technology Centers).

Theoretical Foundations

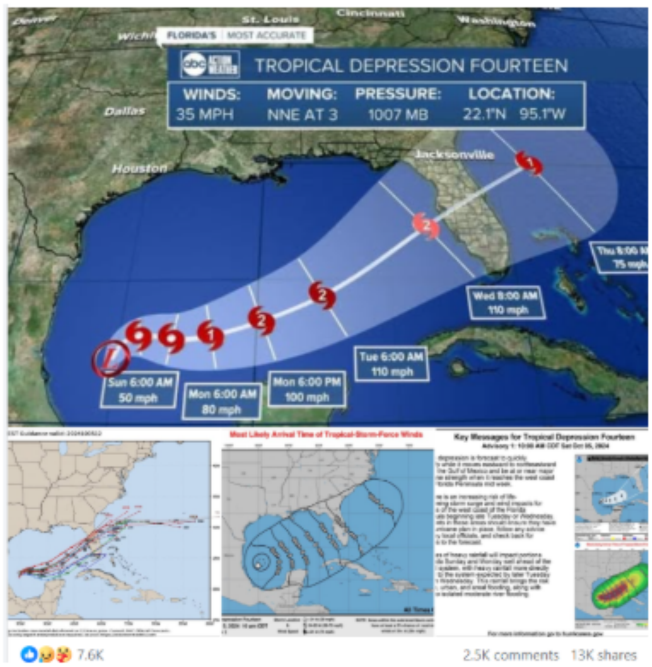
1. In the age of the Internet and social media, the public needs access to both the scientific process and its products. Scientific information is generated by the scientific community using theories and methodologies established within their domains [36]. Traditionally, much of this information has reached the general public through media, as well as via government, professional, or community organizations [37]. These groups have served as gatekeepers, selecting and distributing information for public consumption [38]. In recent years, however, this landscape has shifted dramatically. Members of the public increasingly access and communicate scientific information through the Internet and social media, leading to a proliferation of misinformation, misunderstanding, and misinterpretation of scientific findings [39]. To navigate this complex information ecosystem, scientifically literate citizens must be able not only to interpret scientific information themselves but also to distinguish credible scientific information from misinformation [40]. Achieving this requires an understanding of epistemics—the processes by which scientific information is generated, validated, and communicated [15]. Such epistemic understanding enables the public to act as critical consumers of scientific information and make informed decisions while navigating personal and community experiences amidst an overwhelming amount of information with varying levels of scientific credibility [40]. APLUS will develop an intuitive simulation experience generating storm tracks similar to those produced by scientists and communicated by weather authorities. We will test whether the simulation experience enhances public Uncertainty Literacy in Atlantic Forecasting (ULAF) defined as interpretation of storm track visualizations, attribution of uncertainties in the visualizations to the simulation process, and adequate perception of risks.

2. Scientific uncertainty in the scientific process should be part of hazard literacy. Uncertainty is an inherent aspect of science [41], particularly in the study of hazards and risks [42]. Hazards can have profound personal, societal, economic, and environmental impacts [43], making it critical to effectively communicate hazard forecasting information, including the uncertainties embedded in these forecasts, to the general public [44]. However, identifying effective methods to communicate uncertainty in hazard scenarios remains a persistent challenge for scientific authorities, who must navigate the delicate balance of maintaining public trust in scientific information while supporting informed decision making [13]. Uncertainty is often conveyed in technical terms such as probabilities, error margins, and confidence intervals or through visual tools such as the cone of uncertainty used in storm forecasting [45]. Research shows that the public put a greater trust in scientific information when uncertainty is presented in technical terms [46], [47] as compared to when it is presented as controversies [48], gaps in knowledge [49], or methodological limitations [50].

Figure 1. An example of a local meteorologist's hurricane risk communication on Facebook.

XXX YYY, October 5th facebook post
Chief Meteorologist at ZZZ, Tampa Bay
(1.2M followers) (2.5K comments; 13K shares)
Saturday afternoon Update:

1. The ENTIRE west coast of Florida is in the projected path of TD#14. Reasonably though, the highest likelihood of landfall is between Hernando County and South Florida.
2. This is an extremely unusual track. Instead of coming in from the coast, Milton will come in from the West. That makes the exact point of landfall extremely important. Areas just to the SOUTH of landfall will see the worst of the surge. Areas North will see an offshore flow and the water will be pushed away.
3. Same with wind. Remember the highest winds with any storm are usually about 5 to 10 miles from the eye. Outside of that zone, the winds are much lower. Winds will pick up later Tuesday and landfall will be Wednesday.
4. Models show a track pretty much from the Bay Area to Naples. Anyone in that area could see landfall. Right now, the NHC says Cat 2 on the fringe of 3. Scary numbers, but again, those numbers will only be in an area around 10 miles from where the eye makes landfall. Winds will be much lower outside the point of landfall
5. Now is the time to prepare. We have 4 days to do what you need to do to keep you and your family safe. The margin of error in this forecast is about 150 miles, but the reality is, someone on the West coast is gonna see this make landfall. Don't underestimate it. Impacts outside that 10 mile area will probably be 75-85 mph winds about 35 miles from the center. 55-75 outside of that. Intensity isn't the big question, exact track is.



Risk communication has evolved from definitive and clear-cut messaging to approaches that incorporate probabilistic information, addressing the complexities of decision-making under uncertainty [51]. Communicating scientific uncertainty in weather forecasting has become a routine practice. For example, the National Hurricane Center has been releasing cone of uncertainty graphics to the public since 2002, which were quickly adopted by weather forecasters nationwide [10]. More recently, spaghetti plots have gained prominence, circulating widely through social media platforms [5]. In response to this shift, APLUS conceptualizes ULAF in three facets. The first area involves interpreting storm visualizations, such as the cone of uncertainty and spaghetti plots. Figure 1 illustrates how a local meteorologist utilized social media (Facebook) to communicate storm information during Hurricane Milton in the fall of 2024. The post featured probabilistic visualizations accompanied by the meteorologist's interpretation, highlighting where the storm might be headed, when it might arrive, and its potential impacts, all while emphasizing the uncertainties inherent in these forecasts. The meteorologist, with a substantial following of 1.2 million, garnered significant engagement, with the post receiving 2,500 comments from the public. While these probabilistic visualizations have become widely used, research into public understanding of cones of uncertainty reveals frequent misconceptions, including misinterpretations of depicted storms' potential future locations and associated risks [10], [11]. Spaghetti plots appear to help the public grasp the expanding areas potentially threatened by a storm, but they also lead to an overemphasis on localizing risks to specific real-world locations, which can mis-guide decision making [7]. The second facet of ULAF focuses on the origin of uncertainties portrayed in these storm visualizations, which are rooted in the simulation-based modeling process. The third facet addresses risk perceptions derived from these visualizations. Uncertainty plays a pivotal role in shaping public trust in forecasting information and influences how individuals perceive storm-related hazards and make preparedness decisions [52].

3. Computer simulations provide a unique learning opportunity to explore storm forecasting and uncertainties. In investigating natural phenomena, scientists rely on observations that generate data and models they can manipulate [53]. These models range from concrete and tangible to abstract and algorithmic, incorporating symbols, equations, and mathematical or computational operations [54]. By using models, scientists can isolate and control specific aspects of a phenomenon to devise explanations, refine theories, and make predictions. Each model, developed within a unique material and conceptual context, aims to capture the critical elements of a system believed to influence its behavior and outcomes [55]. Simulations are computational models defined by mathematical constructs and operations that allow manipulation of initial parameters [18]. Simulations illustrate theories by representing relationships between system variables and external factors surrounding the system and are useful when manual calculations are impractical due to complexity or time constraints [56]. Simulations rely on computational algorithms to perform stepwise calculations, iteratively refining parameter estimations of a system that extends beyond the initial state of the system [17]. Simulations used in science education are interactive visualizations that enable learners to easily manipulate variables and interpret results in an engaging and user-friendly manner [57]. Simulations are commonly used in formal science classrooms, effectively teaching scientific concepts, processes, and the nature of science [58], [59] to learners ranging from elementary school to university levels [60]. Their visual and interactive nature also makes them particularly beneficial for English Language Learners as they help compensate for language barriers [61]. Simulations can also be used to help learners consider the uncertainties involved in extending simulation-based evidence to real-world phenomena [27].

Weather forecasting involves using numerical models to simulate the evolution of the atmosphere and its interactions with other components of the Earth system [62]. These models rely on equations describing atmospheric flow and are highly sensitive to the initial conditions provided. This sensitivity leads to the growth of forecast errors over time, ultimately limiting predictability [63]. Two primary sources of uncertainty in weather forecasting come from the simulation process and the simulation outcome. In the simulation process, (1) initialization uncertainty involves the challenge of accurately capturing the true state of the Earth system at the start of the simulation, and (2) model uncertainty involves the approximations, assumptions, and algorithmic limitations to capture complex physical processes. As a result of these modeling uncertainties, simulation outcomes also contain uncertainty defined as ensemble uncertainty, reflecting variability across multiple models or model runs at a given time, and temporal uncertainty, reflecting the run-to-run variability within a single model due to updates with new observational weather data. Since engaging directly with model uncertainty requires specialized domain and technical expertise, APLUS will design the NAS Explorer for the public to experience initialization, ensemble, and temporal uncertainties.

Project Design: Simulation-based Forecasting Experience

APLUS Framework. The solid arrows in Figure 2 illustrate the conventional pathway in which scientific information is created, distributed, and consumed in the context of North Atlantic storm forecasting. Using historical and current data around atmospheric and oceanic conditions together with established theories, scientists develop computational models to simulate storm behavior. When a storm occurs, its data, along with other environmental conditions, are observed and entered into various simulations (e.g., commonly referred to as the European and American models), which generate forecast outputs. A spaghetti plot visualizes these outputs, with each path representing a forecast produced by a different model. Media organizations, government agencies, and community networks distribute forecasting information and visualizations for the general public (Figure 2A). This distribution focuses on presenting the simulation product (such as the spaghetti plots) without conveying the processes behind it [64]. Once the information reaches the public (Figure 2C), it is subject to diverse interpretations shaped by personal knowledge and experiences and sometimes competes with misinformation [65]. In the APLUS project, we will create a parallel pathway for public engagement, enabling individuals to participate in the process of simulation-based storm forecasting (Figure 2B). We anticipate that this experience will enhance users' understanding of uncertainties inherent in the simulation-based forecasting process as practiced by scientists, thereby improving their ability to understand weather authorities' probabilistic storm visualizations and adequately perceive storm risks.

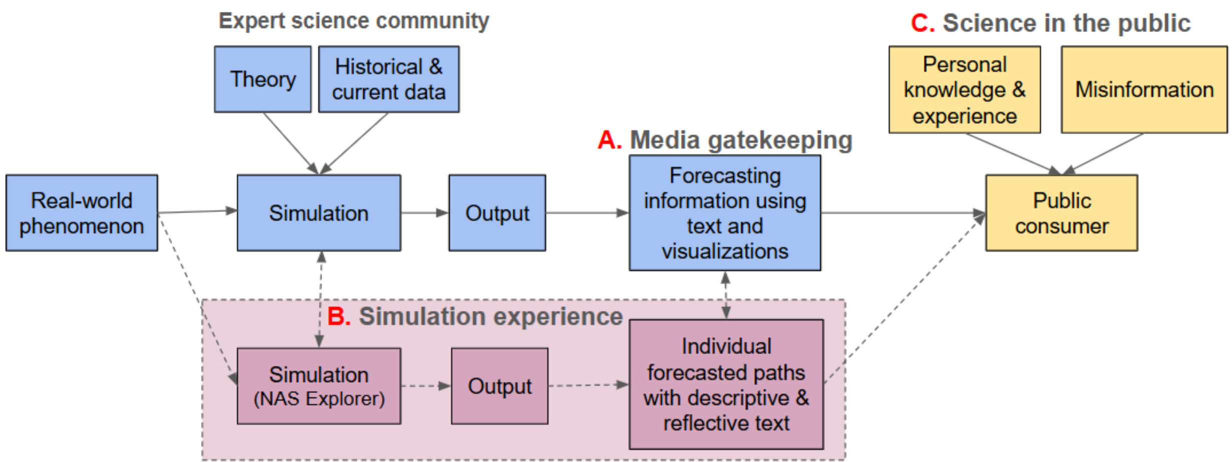


Figure 2. North Atlantic storm forecasting information pathways: Conventional (•—•) vs. APLUS (---►)

Simulation Design. To develop the NAS Explorer, we will leverage an existing simulation interface called Hurricane Explorer developed as part of the GeoHazard project (see Results from Prior NSF Support). The Hurricane Explorer (Figure 3) allows users to manipulate various factors influencing storm behavior, such as average sea-surface temperature and the locations and sizes of atmospheric high and low-pressure systems. Users can observe how these factors affect the storm's track and intensity as it progresses over time, how wind direction and speeds react to the storm's movement, and what storm surge and precipitation amounts it generates.

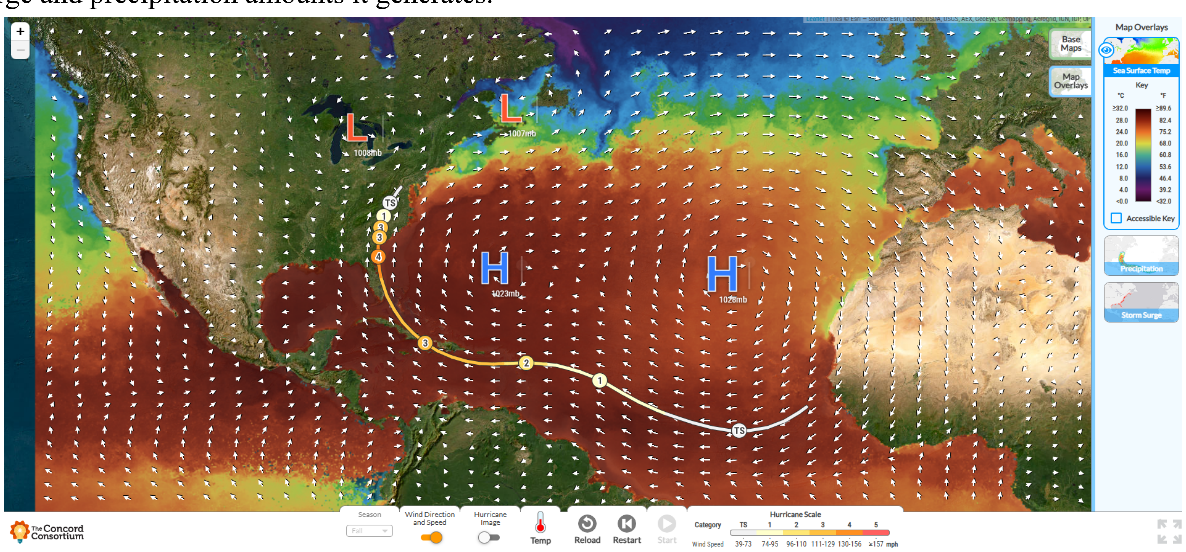


Figure 3. Hurricane Explorer

We will collaborate closely with four scientists—two with hurricane expertise and two with Nor'easter expertise—to modify and enhance the Hurricane Explorer into the NAS Explorer capable of simulating the process of creating forecast tracks for both hurricanes and Nor'easters. The NAS Explorer will be a streamlined and simplified version of scientists' simulation models, which typically require supercomputers. This design will make simulation runs quick for the users to generate storm tracks while preserving the essence of the simulation forecasting process, e.g., incorporating current storm data (e.g., relevant positional data on current storms and atmospheric conditions such as the location and intensities of high- and low-pressure systems), manipulating environmental conditions, and observing storm track responses to these adjustments. As such, the NAS Explorer will not be designed to serve as an accurate forecasting tool, a fact that will be made clear to participants. It will, however, be designed to connect closely enough to currently unfolding storms to spur engagement and provide relevant insights into the process and nuances of storm track forecasting.

As users engage with the NAS Explorer, we anticipate that they will develop an epistemic understanding of the simulation-based storm forecasting process. First, the simulation demonstrates how a storm track responds to interactions between storm-specific factors (e.g., position and initial direction) and environmental factors (e.g., atmospheric high- and low-pressure systems and average seasonal sea surface temperatures). Second, users will encounter *initialization uncertainty*, as the complexity of storms and the limitations of available data make it challenging to accurately reflect storm conditions within the simulation. Third, with different storm and environmental parameter values specified, users will generate varying storm paths, experiencing *ensemble uncertainty*. This variability, reflecting users' choices, produces a set of potential storm tracks, mirroring the ensemble uncertainties found in real-world forecasts. Participants will explore how changes in input conditions lead to the creation of simulated spaghetti-plot visualizations, offering insights into the probabilistic nature of forecasting. Finally, the NAS Explorer will allow users to save and share their ensemble visualizations and re-engage with the simulation as the storm evolves, incorporating updated inputs from real-world forecasts. This iterative use will expose users to *temporal uncertainties*, as they observe differences in their forecasted tracks across sessions driven by evolving real-world storm conditions.

Designing the Simulation-Augmented Forecasting Information (APLUS) Pathway. We will develop the simulation-augmented pathway consisting of three activities: accessing storm information, using the NAS Explorer, and communicating about the storm. By consolidating these activities on a single platform—corresponding to A, B, and C in Figure 2—which typically occur across multiple websites or apps, we streamline user engagement and more effectively isolate the impact of the simulation experience on improving participants' ULAF during the pilot study employing a randomized control trial.

Storm information access (Figure 2A). Weather agencies such as the National Hurricane Center and the National Weather Service make a wealth of visualizations available, releasing them at regular time intervals for storms (every six hours for hurricanes, for example). When designing this activity, we will evaluate various types of storm visualizations to determine the most effective ones for providing participants with live-updated information on any new storms identified by weather authorities. In addition, dedicated sections of the activity page will provide explanatory annotations for relevant features in these visualizations. For the duration of each identified storm, we will include text updates issued by the weather agencies such as advisories, descriptions of impact areas, and recommended actions.

Simulation experience (Figure 2B). Participants will be guided to use the **NAS Explorer** to create their own ensemble storm tracks, take snapshots of their plots, annotate them, and reflect on the factors within the simulation that influenced changes in the storm tracks. Participants will be informed that the **NAS Explorer** is a highly simplified tool and is not designed to produce accurate forecasts.

Communication (Figure 2C). According to Bica et al. [5], it is common for the public to engage actively with storm visualizations by commenting, providing suggestions, exchanging ideas, and asking questions. For example, the Facebook post featured in Figure 1 garnered 2,500 comments. The message board activity in this project is designed to mimic the way scientific information is diffused and discussed publicly, providing an opportunity to understand what participants learn from using the simulation and their grasp of probabilistic storm visualizations. Participants will be encouraged to share ideas, ask questions of other participants or scientists, and discuss their experiences related to storm forecasting. The message board will be monitored and facilitated by AMS staff and scientists as well as CC researchers.

As storms develop during the hurricane and Nor'easter seasons (June to April), participants will be notified via email and social media when storm tracking begins and will have the flexibility to engage with the storm tracking activities as described above, as frequently as they choose, over the course of the storm's progression (typically 5 to 7 days). This simulation-augmented storm information pathway represents our initial design and will be iteratively refined and updated based on participant engagement data and feedback collected during the project's design research phase.

Project Design: Research

RQ1: What is the state of adult participants' Uncertainty Literacy in Atlantic Forecasting (ULAF)?

a. How well do adult participants interpret storm forecast visualizations (e.g., cones of uncertainty and spaghetti plots), attribute sources of uncertainty in the visualizations to the simulation process, and perceive risks conveyed by them?
b. What relationships exist among forecast interpretation, uncertainty attribution, and risk perception?

Research Design. To effectively respond to natural hazards such as hurricanes and Nor'easters, individuals need the ability to interpret and use probabilistic storm forecasting information [66]. We will develop and employ a survey that includes demographic questions and instruments related to three constructs of ULAF, i.e., storm visualization interpretation, uncertainty attribution to the simulation process, and risk perception. The survey design will be informed by prior studies that investigated public understanding of hurricane cones of uncertainty [4], spaghetti plots [7], and risk perception [52]. We will apply the construct modeling approach [67] to validate an instrument and an associated measurement scale for each construct. We will also employ structural equation modeling (SEM) to examine the relationships among the three constructs [68]. The survey will be translated to accommodate participants' language preferences.

Participant Pool. Hurricanes and Nor'easters significantly impact communities along the Atlantic Coast and Gulf of Mexico, spanning 19 states and Puerto Rico, including major population centers such as Houston, Tampa, Miami, Washington, New York, and Boston. To investigate public ULAF, participants in these regions will be recruited through various AMS channels and AMS-affiliated weather-focused social media influencers. AMS will recruit participants through its contact lists of thousands of educators and weather enthusiasts, its Weather Band community engagement program, and its various social media channels. We will also collaborate with social media influencers who work within the weather domain to recruit through their own channels, leveraging AMS's robust relationship with many of these influencers through its new Certified Digital Meteorologist program and its Board on Broadcast Meteorology. During recruitment, we will pose demographic questions such as age, gender, and ZIP codes (socioeconomic status indicator based on their Social Vulnerability Index (SVI) [69], a place-based index, database, and mapping application designed by the CDC to identify and quantify communities experiencing social vulnerability), selecting participants so as to achieve a balanced representation across these demographic variables. We will use this recruitment and selection method for all three years. To establish reliable measurement scales using Rasch analysis for polytomous items, such as those measuring risk perceptions, a sample size of at least 300 respondents is required to ensure stable item properties and person ability estimates [70]. However, to anticipate and account for the variability and potential dropout rates often encountered in informal learning research, we will collect survey data from 400 respondents to ensure robust and reliable findings. We will administer the survey online and provide participants with $25 gift cards as compensation upon survey completion.

The **ULAF Survey** consists of the following sections.

Demographics. Participants will respond anonymously to questions about their demographic background, including gender, age (in years at last birthday), race/ethnicity (using U.S. Census categories), household gross annual income (in dollars), ZIP code, and education level (high school, community college, college, and graduate school). ZIP codes will be used to infer participants' community socioeconomic status by cross-referencing census data and identifying whether they reside in hazard-prone counties. Participants will provide an open-ended response describing their prior experience with hurricanes or Nor'easters.

Construct 1: Storm forecast visualization interpretation. For the hurricane cone of uncertainty questions, we will adapt questions from Evans et al. [4] who surveyed 2,847 Florida residents to assess how accurately they interpret the information conveyed by a hurricane cone of uncertainty visualization. They responded either true or false to questions such as whether the cone of uncertainty graphic shows: the size of the storm (false), the type of storm (e.g., tropical storm, hurricane; true), where the storm could go in the next few days (true), regions where watches and warnings have been issued (true), and areas where damage will occur (false). Residents answered additional questions about increasing uncertainty about the forecast as time progresses (true), whether areas outside the cone are not predicted to be damaged (false), all possible paths of the storm's center (false), and the extent of the storm's damage increasing over time (false). We will adapt these questions as needed to align with Nor'easter cone of uncertainty visualizations, ensuring item consistency between both storm types.

Spaghetti plots explicitly convey forecast uncertainty by illustrating a set of distinct potential storm paths that are simultaneously possible at a given point in time [7]. Although a small body of literature has explored the interpretation of weather visualizations, including spaghetti plots [71], [72], no well-established questions currently exist for assessing understanding of spaghetti plots for hurricanes and Nor'easters. To address this gap, we will develop a set of true/false questions to evaluate participants' ability to interpret key features of spaghetti plot visualizations (e.g., "Different lines represent predicted paths from different simulation runs" [true]) and their understanding of the uncertainty conveyed (e.g., "Although the actual storm track is unknown, the storm is assured to fall within the blue lines shown" [false]). These items will be reviewed by scientists to ensure clarity, accuracy, and alignment with the intended concepts before being incorporated into the online survey.

Construct 2: Uncertainty attribution to simulation process. To address the lack of existing questions assessing how simulations are used to generate storm paths, we will develop a set of true/false questions targeting initialization, ensemble, and temporal uncertainties. For initialization uncertainty, questions will focus on the limitations and approximations inherent in constructing simulation models, such as "Scientists know the environmental conditions at the present storm position" (false). For ensemble uncertainty, questions will address variability across different models, such as "Each line in a spaghetti plot represents a storm path predicted by a different simulation run" (true) or "When runs disagree, it means the storm path cannot be predicted at all" (false). For temporal uncertainty, questions will explore the variability between successive model runs, such as "A forecast made today may differ from one made tomorrow because new data is incorporated into the simulation" (true). These questions will be reviewed by scientists for clarity, accuracy, and domain relevance.

Construct 3: Risk perception. We will utilize items developed by Trumbo et al. [52], who identified and validated cognitive and affective dimensions of risk perception related to hurricanes. The cognitive dimension focuses on the extent to which individuals perceive hurricane risk objectively, such as their belief that the risk of hurricanes is increasing or their perception of how well scientists understand hurricane risk. The affective dimension, by contrast, addresses personal feelings and emotions tied to hurricane risk, such as the level of dread individuals feel about the possibility of a hurricane or how anxious or angry the thought of a hurricane makes them [73]. We will adapt these items for Nor'easters.

Data Analysis. To answer **RQ1a**, we will use descriptive statistics such as frequencies and percentages to summarize participants' performances on individual survey items. This analysis will provide an overview of how well participants interpret storm forecast visualizations (e.g., cone of uncertainty and spaghetti plots), attribute sources of uncertainty to the simulation process, and perceive risks conveyed by these visualizations. To address **RQ1b**, we will establish measurement scales for storm visualization interpretation, uncertainty attribution, and risk perception. Initially, we will employ Classical Test Theory to evaluate the psychometric properties of each instrument, including Cronbach's alpha (to assess inter-item reliability), item discrimination indices, and item-test correlations [74]. These properties will guide the refinement of the instrument by identifying items to modify, add, or eliminate. Next, we will apply Rasch analysis [75] to develop a robust unidimensional measurement scale for each construct of ULAF. Prior to conducting Rasch analysis, we will test for unidimensionality by performing an exploratory factor analysis (EFA). Using a Scree plot, we will determine the number of significant latent factors [76], with eigenvalues of 2.0 or greater considered substantial [77]. We will also assess the local independence assumption using the Q3 statistic [78], ensuring that the mean Q3 statistic is close to $-1/(n-1)$, where n is the number of items in the instrument. Once the Rasch analysis is complete for each construct, we will obtain Rasch ability estimates for participants. These estimates will allow us to compare overall trends across demographic factors using statistical tests such as independent and paired t-tests and one-way ANOVAs on each construct. To further explore relationships among three constructs of ULAF, we will employ Structural Equation Modeling (SEM) if appropriate [68]. SEM can test direct, indirect, and bidirectional relationships among the three constructs, examining how these elements of ULAF interact.

RQ2: How do adult participants interact with and engage in the simulation forecasting experience?

a. What patterns emerge in participants' use of the NAS Explorer to create storm tracks?
b. What uncertainties do participants notice, and how do they communicate about these uncertainties?
c. How do patterns of simulation use and uncertainty communication contribute to ULAF?

Research Design. We will use RQ2 to guide the development of the NAS Explorer and the simulation experience within the context of the storm forecasting information pathway. We will employ the design-based research paradigm [79], [80] where initial design choices are informed by established learning theories, existing research findings, and the experiences of prior designers and iteratively refined based on the design implementation with participants and their feedback. RQ2a seeks to uncover participants' uses of the NAS Explorer in forecasting storm tracks; RQ2b addresses the project's goal of fostering awareness and understanding of uncertainty in storm forecasting, which can be observed in their description of forecasts and during message board communication; and RQ2c connects engagement behaviors with the three constructs of ULAF. For the design research described below, we will recruit participants from the same pool as RQ1.

Simulation design and testing (Year 1). We will develop a prototype of the NAS Explorer and conduct two "clinical" trials using think-aloud methods with 10 adult participants recruited online. Participants will engage through video conferencing in two sessions (5 in each session), allowing for iterative revisions to the simulation. Each participant will be compensated \$50 for one hour of involvement. Testing will focus on usability (e.g., "Was the interface intuitive to navigate and accessible?") and functionality (e.g., "Did the simulation respond as expected to your inputs?"). Feedback will guide refinements to ensure the NAS Explorer is user-friendly and meets its design goals.

Simulation experience design and testing (Year 2). We will recruit a total of 100 participants using the recruitment method outlined in RQ1: 50 participants to follow hurricanes from September to November and another 50 participants to follow Nor'easters from January to April. They will take the ULAF survey as pre-test and post-test and engage with the North Atlantic Storm forecasting for six times, two times per storm for three storms. AMS research staff will monitor storm activity and announce when a storm is approaching the U.S., a period typically lasting five to seven days. The estimated time commitment for participants includes 15 minutes each for the pre-survey and post-survey, 90 minutes for forecasting six times (two per storm, three storms, each lasting 15 minutes), and additional time for interviews. Participants will receive a total compensation of \$150 for their involvement. The goal is to gather data to improve the simulation and the storm forecasting activity and to learn how best to reach and engage participants during a storm event.

Data Source. ***ULAF survey***. Participants will complete the pre-ULAF survey before and the post-ULAF survey after the simulation-based forecasting period. ***Computer logs*** will capture and automatically record a user's interactions across three activity sections. These include reading and viewing weather authorities' forecasting information in the first section, using the simulation in the second section, and writing, reading or responding to posts in the third section. Specific user actions to be recorded include forecasting information accessed and viewed, clickstreams of simulation elements, factor choices made, forecast tracks created, and messages viewed and responded to. All logs will be time-stamped and linked to anonymized user IDs, enabling the integration of interaction data with participants' pre- and post-ULAF surveys for analysis. ***Message board posts***. We will collect and download all messages for analysis. The message board activities may be guided or facilitated by AMS staff and/or scientists to encourage participants to reflect on various aspects of uncertainty in storm forecasting emphasizing initialization, ensemble, and temporal uncertainties. ***Interviews*** will focus on exploring participants' interactions with the NAS Explorer, their thought processes during the activities, and their reflections on the overall experience. We will also prompt participants to talk about the uncertainties they noticed during the simulation process and how they interpreted or addressed those uncertainties. Participants will reflect on their communication experiences, particularly how they shared observations or engaged in discussions about uncertainties on the message board. Interviews will further explore how the simulation-based experience influenced participants' understanding of storm forecasting and solicit feedback on all three activity sections, including any technical difficulties participants encountered and suggestions for improvements to make the forecasting experience more intuitive and engaging.

Data Analysis. To address **RQ2a**, we will first extract relevant log events and assign them to forecasting information access, simulation use, and message board interactions. A descriptive comparison of individual actions will be conducted to identify patterns and trends in each section of the forecasting experience. Using time-stamped data, we will examine the sequence of actions to analyze how participants interact with the simulation, tracking the frequency and duration of these activities. See Table 2. For **RQ2b**, we will perform content analysis of message board posts. Quantitatively, the analysis will focus on metrics such as the frequency of posts addressing specific types of uncertainty, the average word count or complexity of posts related to uncertainty, and the proportions of posts corresponding to coding categories defined in Table 3. Qualitatively, we will assess the depth of participants' engagement with uncertainty topics and the ways they articulate their understanding. To answer **RQ2c**, we will compare several quantitative measures derived from Tables 2 and 3 between participants who demonstrate improvement in ULAF and those who do not. This information will inform what simulation use and message board interaction measures to use as mediators in the pilot study (RQ3). To strengthen the validity of these findings, interviews will serve as a secondary data source for triangulation. Interviews will provide qualitative insights into participants' thought processes, perceptions of uncertainties, and reflections on their engagement with the NAS Explorer and the overall forecasting experience. We will analyze data from interviews thematically, focusing on recurring patterns that align with or diverge from trends observed in log data and message board activities.

Table 2. Computer log analysis plan for each participant

Section 1: Information access	Section 2: Simulation use	Section 3: Message board interactions
● Information viewed (e.g., storm advisories, impact areas, visualizations). ● Frequency of access to specific information.	● Storm initialization settings. ● Parameter adjustments. ● Number of unique simulation runs ● Number of simulation runs. ● Time spent in simulation.	● The number of messages posted. ● Messages read/viewed. ● Specific uncertainties described. ● Comparison between their simulation outputs to professional forecasts.

Table 3. Message board communication analysis plan

Uncertainty noticed	Uncertainty expression	Uncertainty reasoning
● Modeling uncertainty about limitations in storm initialization. ● Ensemble uncertainty about variability across different models or model runs. ● Temporal uncertainty about changing forecasts over time with new data input. ● Implications of uncertainty on decision-making and preparedness.	● Identification of specific sources of uncertainty. ● Indirect expression of uncertainty through questioning the reliability of a forecast or seeking clarification. ● Using evidence such as simulation outputs, professional forecasts, or prior knowledge.	● Simple observations. ● Detailed discussions of causes or implications of uncertainty. ● Misconceptions about uncertainty.

RQ3: What impact does the simulation-augmented forecasting experience have on adults' ULAF when encountering weather authorities' forecasting information?

a. Does engaging with the simulation-augmented forecasting experience improve adults' ULAF?
b. Do the effects differ by demographic characteristics, such as gender, race, storm experience, socioeconomic status (SVI), and education level?
c. Do the effects depend on the extent of participants' engagement with the simulation-augmented forecasting experience?

Research design. We will conduct a randomized control trial (RCT). From the participant pool described in RQ1, 300 adult participants will be recruited, ensuring balanced representation across demographic factors. Participants will be randomly assigned to one of two pathways: the APLUS pathway (treatment), refined in Year 2, which includes storm forecasting information access, simulation use, and message board communication, or the conventional pathway (control), which includes only information access and message board communication. Both will be featured on the same platform but separately managed.

Power Analysis. We conducted a power analysis using SPSS with a linear regression model. A sample size of 200 participants is sufficient to achieve a statistical power of 0.90, assuming a significance level of 0.05, a two-tailed test, and a medium effect size (partial correlation of 0.30) with a total of up to 8 predictor variables. Due to the expected added noise associated with informal education settings, we are increasing our sample size to 300.

Participants and Data Collection. Participants in both conditions will complete a ULAF survey (see RQ1) before and after the study. To analyze engagement patterns, we will collect computer logs of participant interactions and message board postings throughout the study. Each participant will receive a minimum of $100 as compensation for their participation, which includes completing a pre-survey, engaging in storm forecasting activities for two storms, and completing a post-survey. Additional storm tracking will be encouraged and compensated at $25 per extra storm, with total compensation potentially increasing to $200 for participation in six storms. Each storm will be tracked for approximately 5-7 days, depending on its development, allowing participants the flexibility to visit multiple times for the same storm or to explore different storms. The variation in storm tracking frequency will be critical in assessing the impact of the simulation-based forecasting activity and will serve as a measure of intervention dosage.

Data Analysis. We will conduct the impact analysis using a linear regression method, with the Rasch estimates of three ULAF constructs as dependent variables. For each dependent variable, the regression model will include the treatment variable (coded as 1 for treatment and 0 for control), the pre-survey measure of the dependent variable, and demographic variables such as gender, race, socioeconomic status, storm experience, education level, and residential location. In addition to assessing the overall treatment impact, we will conduct a moderator or subgroup analysis to explore how the treatment effects vary across different demographic groups. This analysis will include an interaction term between the treatment indicator and subgroup variables. For the treatment group, we will undertake a mediator analysis to examine how participants' engagement with the simulation and message boards influences their outcomes. Mediators will include measures identified as significantly correlated with pre-post survey gains in ULAF in RQ2. These measures could be the time spent on simulations, the number of simulation trials completed, the frequency of message board posts, and so on. This analysis will use either a regression model or a structural equation modeling (SEM) approach to explore the relationships between these engagement metrics and learning outcomes [81].

Communication Plan

The target population for the NAS Explorer-based simulation experience includes adults (18 years or older) seeking storm forecasting information online, particularly those living in communities vulnerable to the impacts of hurricanes or Nor'easters. The project also aims to engage stakeholders involved in the communication of scientific information, including scientists, professional weather agencies, and public outreach organizations. For dissemination, CC and AMS will create designated webpages for the APLUS project, featuring the NAS Explorer, associated forecasting activities, research instruments, and findings. For broad dissemination, the project will leverage multiple social media channels, including AMS Facebook pages—AMS Facebook (49K followers) and AMS Education Facebook (1.4K)—as well as the Concord Consortium's Facebook page (2.2K). AMS also publishes *Weather with a Twist*, a Substack newsletter with over 4,000 subscribers that describes meteorological research results in a lighthearted but scientifically rigorous manner. The @*Concord* newsletter, a biannual publication emailed to more than 65,000 subscribers, will feature updates on the project. In addition, the Bulletin of the American Meteorological Society, a magazine with 12,000 subscribers, many of whom are scientists and educators, will highlight project progress and findings. The CC blog will provide updates on simulation availability, learning experiences, research findings, conference presentations, and publications. AMS will also promote the NAS Explorer to its network of nearly 1,000 K-12 educator alumni and include it in their annual meteorology K-12 teacher PD courses (N=~90). The project will disseminate findings widely at research conferences in education, including the American Educational Research Association and the

National Association for Research in Science Teaching, informal science education such as the Association of Science and Technology Centers, and metrology such as Annual Meeting of the American Meteorological Society and International Meteorology Organization biannual Congress. The project will publish findings in academic journals in science education (*Journal of Research in Science Teaching* and *International Journal of Science Education-Informal Science*), educational technology *(Journal of Science Education and Technology*), and meteorology (*Bulletin of the American Meteorological Society*).

Evaluation

The Advisory Board (AB) will provide critical evaluation and feedback. Beginning in Year 1, the AB will actively monitor the project, receiving monthly email updates on progress throughout the project duration. This ongoing communication will facilitate two-way interaction between project partners and AB members, enabling timely responses to questions, insights, and suggestions from both sides. The AB and project partners will meet in person at the end of Year 1 and virtually at the end of Years 2 and 3. These meetings will focus on reviewing the year's outcomes, providing targeted feedback based on members' expertise, and drafting individual reflections and evaluations. Dr. DeWitt, as chair of the AB, will consolidate these contributions to align them with project objectives and milestones. The consolidated and individual evaluations will be shared with the project team and included in the NSF annual report.

Jennifer DeWitt, Ph.D., a Senior Research Fellow in the Department of Education, Practice, and Society at UCL Institute of Education, UK, specializes in fostering science capital through equity-focused research in informal science environments. With extensive experience in evaluating informal science projects, she will ensure the APLUS project effectively meets its informal science learning goals.

Ying-Chih Chen, Ph.D., an Associate Professor of Science Education in the Teachers College at Arizona State University, specializes in engaging learners with scientific uncertainties through modeling practices. With expertise in instrument design and advanced statistical analyses, he will evaluate the development of the ULAF survey and provide guidance on the use of appropriate statistical methods.

Jody Clarke-Midura, Ed.D., an Associate Professor of Instructional Technology and Learning Sciences in the College of Education at Utah State University, promotes inclusivity in STEM education through technology and developing assessment methods utilizing learners' lived experiences. She will advise on the simulation-based forecasting experience design to enhance engagement and inclusivity.

Joseph Trujillo-Falcón, Ph.D., holds a joint appointment at the University of Illinois Urbana-Champaign Department of Climate, Meteorology, and Atmospheric Sciences and the Department of Communication, transitioning to a tenure-track faculty position in 2025. With expertise in hazardous weather communication across multiple languages, he will advise on research methodology, language translation, and improving the communication of storm forecasting information for diverse audiences.

Melissa Bica, Ph.D., a computer scientist and user experience researcher at Workday, studied the diffusion of forecast and risk visuals on social media during hurricanes and how people make sense of representations of risk and uncertainty. She will provide insights, viewed from the industry perspective, into the public's interpretation of forecast visualizations and their engagement with risk communication.

Chad Kauffman, Ph.D., a Professor in the Department of Biology, Earth, and Environmental Sciences at PennWest University, specializes in meteorology education and curriculum development. As an education and outreach consultant for the AMS Education Program, he will evaluate the relevance of the simulation-based forecasting experience in enhancing public understanding of forecasting uncertainties.

Jack Langefeld is a weather forecaster and influencer. As the co-founder/lead forecaster of *Your Chicago Weather*, he provides weather forecasts and storm chasing content to over 40,000 followers through newsletters, social media, and blog posts with a focus of science over hype. He will offer insights on how to engage with the public, other weather influencers, and the social media weather enterprise.

Scientist Panel. Four scientists—two specializing in hurricanes and two in Nor'easters—selected from the AMS membership, will provide expert scientific guidance for the development and formative evaluation of the NAS Explorer. They will share the computational forecasting models used in their research, offer insights on how to simplify these models to maintain scientific accuracy, and review the items for storm forecasting visualization interpretation, uncertainty attribution to the simulation process, and risk perception items featured in the ULAF survey.

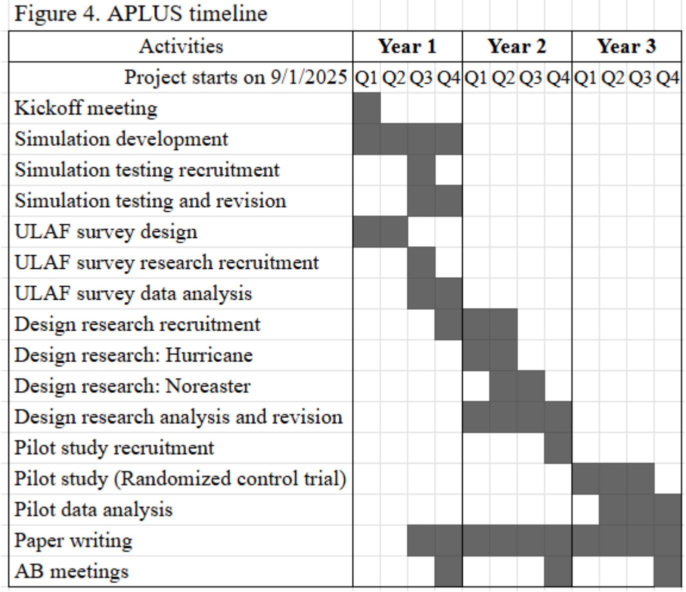
Broader Impacts

This project will create a simulation-based experience to enhance public literacy about scientific uncertainty in hurricane and Nor'easter forecasts. Through engagement with simulations, participants will explore how forecasts are produced and gain an understanding of the uncertainties that emerge from the simulation process. This experience will better equip them to interpret and respond to probabilistic visualizations, such as cone of uncertainty graphics and spaghetti plots, which are crucial tools in storm risk communication. Research findings will be disseminated via journal publications, conference presentations, newsletters, and social media, contributing to advancements in hazard communication, public policy, and disaster resilience planning. APLUS prioritizes inclusivity by engaging underrepresented groups in storm-prone areas, ensuring equitable access to science education and safety resources. The project will make the NAS Explorer, research instruments, and simulation-augmented forecasting experience freely available to informal educators, researchers, scientists, weather authorities, and the public. APLUS will serve as an example of interdisciplinary collaboration among experts in meteorology, science communication, learning sciences, informal science education, and public outreach.

Project Management

Project Timeline. Figure 4 outlines the project milestones and timeline. Project partners from CC, AMS, and PF along with AMS scientists will convene for an in-person kickoff meeting at the start of the project and hold bi-monthly meetings thereafter to review progress, address challenges, and make adjustments. CC will lead the design and development of research instruments, the NAS Explorer, and the simulation experience as well as the analysis of all data, except for computer logs, which will primarily be analyzed by PF. AMS will be responsible for participant recruitment, advising on the scientific aspects of simulation and experience development, and running daily forecasting sessions during storm events. All partners will collaborate on the interpretation of the results and may conduct separate analyses to address specific interests.

Figure 4. APLUS timeline



Project Expertise. Project PI and Co-PIs' expertise and responsibilities are described below.

Hee-Sun Lee, Ph.D. (PI), a Senior Research Scientist at the Concord Consortium (CC), will lead the development and validation of ULAF constructs, design-based research, and the randomized control trial. She will oversee the budget and the timely completion of research and development plans across all partner organizations. Her expertise includes assessment development and validation, epistemic science practices using simulations, the role of uncertainty in modeling, and learning analytics.

Amy Pallant, M. Ed. (Co-PI), a Senior Research Scientist at CC, will oversee the development of the NAS Explorer, the simulation forecasting experience, and the analysis of simulation experience data. She has directed numerous NSF-funded, high-impact Earth and environmental science projects, producing evidence-based digital curricula that integrate computer simulations and data visualizations.

Aaron Price, Ph.D. (Co-PI), Director of Education at AMS, leads a team at AMS conducting online and hybrid K-12 science teacher PD programs and has experience in experimental informal learning research. He will work on scientific content, research methodology, analysis, and dissemination.

Gey-Hong Gweon, Ph.D. (Co-PI), is the CEO of Physics Front (PF), a small business specializing in data analytics for educational applications. He will lead the analysis of computer logs based on his expertise in developing and applying advanced algorithms for tracking knowledge development and identifying patterns in simulation use.