How Undergraduate Students Learn Atmospheric Science

Characterizing the Current Body of Research

Peggy McNeal, Wendilyn Flynn, Cody Kirkpatrick, Dawn Kopacz, Daphne LaDue, and Lindsay C. Maudlin

ABSTRACT: Educators can enrich their teaching with best practices, share resources, and contribute to the growing atmospheric science education research community by reading and participating in the scholarship of teaching and learning in atmospheric science. This body of scholarship has grown, particularly over the past 15 years, and is now a sizable literature base that documents and exemplifies numerous teaching innovations in undergraduate atmospheric science education. This literature base benefits the entire atmospheric science community because graduates of atmospheric science programs are better prepared to enter the workforce. This literature base has not yet been examined, however, to see how well the evidence supports education practices in the atmospheric science education literature. In this study, we characterized that evidence and show that the majority of papers we reviewed share education innovations with anecdotal or correlational evidence of effectiveness. While providing useful practitioner knowledge and preliminary evidence of the effectiveness of numerous innovative teaching practices, opportunities exist for increasing readers' confidence that the innovations caused the learning gains. Additional studies would also help move conclusions toward generalizability across academic institutions and student populations. We make recommendations for advancing atmospheric science education research and encourage atmospheric science educators to actively use the growing body of education literature as well as contribute to advancing atmospheric science education research.

KEYWORD: Education

<https://doi.org/10.1175/BAMS-D-20-0023.1>

Corresponding author: Peggy McNeal, pmcneal@towson.edu

In final form 15 July 2021

©2022 American Meteorological Society

For information regarding reuse of this content and general copyright information, consult the [AMS Copyright Policy](#).

AFFILIATIONS: McNeal—Towson University, Towson, Maryland; Flynn—University of Northern Colorado, Greeley, Colorado; Kirkpatrick—Indiana University Bloomington, Bloomington, Indiana; Kopacz—University of Nebraska, Lincoln, Nebraska; LaDue—University of Oklahoma, Norman, Oklahoma; Maudlin*—Auburn University, Auburn, Alabama

* **CURRENT AFFILIATION:** Maudlin—Iowa State University, Ames, Iowa

Nearly 15 years ago, Charlevoix (2008) charged the atmospheric science community with increasing research into teaching and learning in atmospheric science. Charlevoix noted that atmospheric science as a field becomes more robust and effective at meeting society's challenges when graduates are better prepared. She stated that improvements to education happen more efficiently when university faculty, lecturers, graduate teaching assistants, and others reflect upon and share teaching practices they find to be particularly effective. Charlevoix argued that sharing of effective teaching should be considered a type of scholarship that systematically produces new knowledge in reproducible ways. She asserted that atmospheric science would greatly benefit from developing a body of literature from which educators could draw, saving time and pooling resources, to grow and improve our field.

The atmospheric science community has amassed a sizable education publication record spread among journals such as the *Journal of Geoscience Education (JGE)*, the *Journal of College Science Teaching*, and the *Bulletin of the American Meteorological Society (BAMS)*, with many more publications in journals outside of atmospheric science. However, to advance the type of scholarship advocated by Charlevoix, it is important to characterize the evidence base that supports this work and overall community claims. Our vision for this paper is to connect those who teach atmospheric science with existing education literature and make transparent the strengths and limitations of the research that inform this body of work. In this way, atmospheric science educators might enrich their teaching with evidence-based practice, discover how to effectively share their own work, and participate in the growing atmospheric science education research community.

Charlevoix specifically advocated for scholarship focusing on teaching innovations that address learning goals in undergraduate atmospheric science education. This type of work is known as scholarship of teaching and learning (SoTL) [National Association of Geoscience Teachers (NAGT); NAGT 2020], and is the kind of research we addressed in this study. (For readers interested in how education researchers distinguish between SoTL and other education research, please see the sidebar.) In education research, as in other fields, literature reviews are one method for synthesizing and evaluating ongoing work, identifying gaps, informing new research, and defining norms and standards of a research community. Literature reviews can be an effective method for analyzing multiple studies to assess the evidence supporting community claims, such as curricular interventions used in education (Scherer et al. 2019). We pursued the latter objective in this study. Our research question was, “What is the current state of atmospheric science education research?”

Methods

This project grew among a cross section of atmospheric science and education researchers with a strong interest in atmospheric science education research and a collective understanding of the current literature in our field. Three authors are founding members of the Atmospheric Science Education Research Community, a working group that convened in 2016 to create a vital education research community in atmospheric science. This group conducted the “Involvement in and Perception of Atmospheric Science Education Research Survey” in 2018 (Kopacz et al. 2021). Study participants who had published atmospheric science education papers ($N = 45$) indicated their work had been published in 42 distinct journals, which gave us an idea of the scope of the publication distribution. Five of our authors served on the AMS

DBER and SoTL

Understanding where education research falls along the continuum between discipline-based education research (DBER) and SoTL is important for defining research and interpreting results.

DBER uses content knowledge and cognitive science to investigate best practices for teaching and learning in a discipline (National Research Council 2012). Several disciplines support a broad base of theory-driven research and sustain journals dedicated to the publication of research on teaching and learning within their discipline (Henderson et al. 2017). Most disciplines clearly distinguish DBER as research that tests theory and produces generalizable findings focused on teaching, learning, and ways of thinking, including investigations into the development and nature of expertise and strategies for increasing diversity and inclusiveness within a discipline (NAGT 2020).

On the other hand, the goal of SoTL is “to improve one’s own teaching practice through innovations in pedagogy and

curriculum and to serve as a model for others” (NAGT 2020). Researchers (often course instructors conducting research on their own teaching and students’ learning) are encouraged to gather data that lead to self-reflection, improved teaching practices, and improved student learning. Although SoTL can be course specific, conclusions should be supported by evidence and have broader applications so as to serve as a potential model for other instructors and at other institutions (NAGT 2020). Findings from SoTL studies are sometimes disseminated through workshops or conferences (National Research Council 2012; Kern et al. 2015).

DBER and SoTL studies are both important to advancing an understanding of best practices in atmospheric science education. Overall, the goal is to strengthen the rigor and visibility of atmospheric science education research to ultimately benefit the field of atmospheric science.

Ad Hoc Committee on Atmospheric Science Education Research in 2019. To gain a sense of the publication history of the community, they conducted an informal survey of the literature and identified 43 atmospheric science education research papers. All authors have attended and participated in the Conference on Education held at AMS Annual Meetings and therefore have a broad understanding of the breadth of the existing atmospheric science education efforts and resulting literature. We informally identified 20 additional papers, bringing the total to 63. These were identified through several independent efforts that were not designed to be reproducible; thus, we sought to verify and extend our search through the consistent search method described below.

The search for atmospheric science SoTL papers. We drew upon related work in the geosciences (Bitting et al. 2018; Scherer et al. 2019) to search for relevant literature for our study. We selected three library databases (ERIC, Academic Search Ultimate, and PsycInfo) and five search terms (undergraduate, education, meteorology, atmospheric science, and research) to conduct a search of the literature. To test our search process, each author was assigned one database and used the five search terms individually and with different Boolean operators to produce individual lists of papers that were compiled into one list. We then examined the compiled list to confirm that our process was “catching” what we knew of the publication history and distribution. Because we had previously identified papers and journals that we did not discover through these systematic searches, we added a fourth database (Web of Science) and three more search terms (college, cognition, and learning) and conducted the search again. At this point, the process was successfully identifying the breadth of the literature previously identified and, in many cases, was duplicating results across databases. We ultimately identified a total of 173 papers primarily from five different journals (Table 1).

At this point, we established specific inclusion and exclusion criteria for identifying a set of papers to study (Table 2). We began with the call to embrace SoTL made by Charlevoix (2008). This call applied specifically to increasing SoTL efforts; thus, our inclusion criteria centered on studies focused on teaching innovations addressing learning goals in undergraduate education, including curriculum development. Recognizing that learning gains are not the only constructs worthy of study (LaDue et al. 2022), and that instructors often “wish to contribute to the knowledge base of teaching and learning” (Charlevoix 2008, p. 1659), we retained papers with the intent to improve instruction and learning. Studies that addressed self-efficacy,

Table 1. Sources of papers initially identified and after inclusion/exclusion criteria were applied.

Journal	Number of papers initially identified	Number of papers in the final list
<i>Bulletin of the American Meteorological Society (BAMS)</i>	55	21
<i>Journal of Geoscience Education (JGE)</i>	35	8
<i>Weather, Climate, and Society</i>	27	0
<i>Journal of College Science Teaching</i>	8	5
<i>In the Trenches</i>	2	2
Others	46	11
Total	173	47

motivation, and student willingness to engage in novel instructional strategies (e.g., a flipped classroom) were included, but papers on diversity, expert–novice differences, identifying (but not overcoming) misconceptions, and program development/history were excluded. We chose to focus solely on atmospheric science education research, thus excluding climate change and environmental education papers that did not address atmospheric science concepts.

Finally, in order to focus primarily on the evolution of atmospheric science education research after the Charlevoix call (2008), but to provide a sufficiently lengthy time period upon which to draw conclusions (15 years), we agreed to include papers published in or after 2005. We excluded conference presentations, abstracts, and preprints to confine our search to work published in publications with peer-review or editor oversight. Overall, we came to a consensus on the criteria through group discussions.

We methodically applied our inclusion–exclusion criteria to the 173 papers. Each of the authors selected at least 2 papers to include in a pilot list of 14 papers that we independently determined to include or exclude according to the criteria. We met to compare individual determinations, worked to resolve differences, and further refined our criteria through group discussion. This process was repeated with an additional 12 papers. At this point, we concurred that we had a mutually agreed upon understanding of the application of the inclusion–exclusion criteria. We divided the remaining 147 papers into six sets, and assigned each set to a “primary” and “secondary” (randomly assigned) reviewer who determined whether to include or exclude each paper. In this way, the final determinations to include or exclude, based on agreement between two reviewers, were made on 83% of the papers. We met three times to review and discuss the remaining papers until an agreement was reached by all authors for each paper. This process resulted in 47 papers for evaluation. The 47 SoTL papers are provided in Table 3.

Table 2. Inclusion–exclusion criteria.

Inclusion	Exclusion
SoTL studies	Diversity/inclusion papers
Curriculum development	Expert–novice studies
Is intentional and reflective about improving instruction and learning in atmospheric science	Program development/history
Peer reviewed or has editor oversight	Climate change and environmental education literature
Targets undergraduate instruction (majors or nonmajors), including classroom, field, REUs, and internships	Meeting reports/conference preprints/dissertations
Published 2005–20	Papers focused on misconceptions research
	DBER studies of spatial thinking or other cognitive aspects of learning
	Development of instruments to assess education outcomes

Table 3. List of papers included in this study.

Practitioner wisdom/expert opinion	1	Billings et al. (2019), Clements and Oliphant (2014), Cohen et al. (2018), Coleman and Mitchell (2014), Croft and Ha (2014), Kelsey et al. (2015), Kopacz (2017), Kovacs (2017), Lanicci (2012), Lerach and Yestness (2019), Mandrikas et al. (2017a), Morss and Zhang (2008), Nelson et al. (2019), Shapiro et al. (2009)
Qualitative and quantitative case studies	2A	Davenport (2018), Drossman et al. (2011), Godfrey et al. (2011), Kahl (2017), Kahl and Ceron (2014), Laird and Metz (2020), Neves et al. (2013), Roebber (2005), Schultz et al. (2013), Shellito (2020), Yow (2014)
	2B	Barrett and Woods (2012), Bond and Mass (2009), Cervato et al. (2011), Charlton-Perez (2013), Collins et al. (2016), Hepworth et al. (2019), Hopper et al. (2013), Kahl (2008), Mandrikas et al. (2017b), Mandrikas et al. (2018), Mullendore and Tilley (2014), Quardokus et al. (2012), Roberts et al. (2010), Tanamachi et al. (2020)
	2C	Cobbett et al. (2014), Grundstein et al. (2011), Harris and Gold (2018), Horel et al. (2013), Illari et al. (2009), Jeong et al. (2019), Market (2006)
Qualitative and quantitative cohort studies	3	Mackin et al. (2012)

Adapting and applying the Geoscience Education Research Strength of Evidence Pyramid.

To address our research question, we set out next to evaluate the strength of evidence supporting conclusions described in the papers. “Strength of evidence” refers to *whether* findings were substantiated, the lack of which could indicate researcher bias in the claims made. It also refers to *how* findings were substantiated, use of custom-designed instruments and surveys, or how nondiverse study populations limit generalizability of findings. In general, we think of “strength of evidence” as how well a study demonstrates that an education practice was the cause of changes observed in the students.

We started with the Geoscience Education Research Strength of Evidence Pyramid (St. John and McNeal 2017; Fig. 1). Consisting of five levels for characterizing claims based on the strength of the evidence presented, the Geoscience Education Research Strength of Evidence Pyramid can be used as a rubric to evaluate the strength of evidence of generalizable claims in education research efforts. The shape—a pyramid—represents the relative number of papers at each level. Specifically, the large number of papers that support the foundation of the pyramid describe studies that support a smaller amount of research at the top. The higher levels of the pyramid do not necessarily identify better, stronger, or more rigorous studies. Instead, studies at higher levels of the pyramid have increased generalizability of their conclusions or claims, based on the strength of evidence presented and the research design.

The first level of the Geoscience Education Research Strength of Evidence Pyramid, Practitioner Wisdom/Expert Opinion, describes wisdom about teaching and learning. While containing unsubstantiated conclusions, these papers are valuable because they inform the community by highlighting potential education challenges and promising education practices

Fig. 1. The Geoscience Strength of Evidence Pyramid (St. John and McNeal 2017).

(St. John and McNeal 2017). Work at this level can involve peer review, but there is often no or editor-only oversight. The second and third levels of the pyramid, Quantitative and Qualitative Case and Cohort studies, introduce measures of evaluation and systematic analysis of that evidence to show effectiveness of the education practices described. The studies with the strongest evidence at these levels use validated instruments and recognized education research methodologies to study a class or course. By investigating a broader cohort such as a cross section of courses, institutions, and/or populations, researchers can further reduce bias, increase sample size, and strengthen generalizability and evidence for claims.

The fourth and fifth levels of the pyramid, Meta-Analyses and Systematic Reviews, build on studies forming the foundation of the pyramid. They consolidate the results of multiple investigations, generalize findings for broader populations, elevate confidence in community claims, recognize trends, and identify gaps. Meta-Analyses involve application of statistical methods to look at a broad suite of existing quantitative or qualitative data. Systematic Reviews use transparent methods to identify, select, and evaluate relevant published literature on a particular topic (St. John and McNeal 2017).

Our list of 47 papers was sufficiently small that we could discuss all of the papers as a group over multiple meetings. Before each meeting, each group member read the same subset of papers and preassigned each paper to a level on the Geoscience Education Research Strength of Evidence Pyramid. In addition to the salient criteria found in the Geoscience Education Research Strength of Evidence Pyramid, we also heavily referenced the paper in which the pyramid was published (St. John and McNeal 2017), which contains additional discussion and motivation for each level. To increase the reliability of our process, we asked Karen McNeal, an author of the St. John and McNeal (2017) paper, to read and preassign a strength of evidence level to a subset of papers. The author joined one meeting where these papers were discussed, characterized, and compared to her preassignment. As a result of her interpretation, our discussion at that meeting, and further discussions in subsequent meetings, we divided the Quantitative and Qualitative Case Studies category (Level 2) first into two, and subsequently three, sublevels to further distinguish the large amount and wide range of papers represented within Level 2. Second and third rounds of review ensued in order to recharacterize the papers using the further delineated rubric, which is shown in Fig. 2.

At each of our meetings and through the process of characterization, we identified and discussed the methods, strengths, weaknesses, and how well findings were conveyed in individual studies, along with common practices and gaps in the papers we reviewed. Extensive memo writing, both by individual members and as a collective summary of each meeting, documented these observations.

Representative papers

Our final list of characterized papers is shown in Table 3. These papers exemplify atmospheric science education research at each level of

Fig. 2. The rubric developed to evaluate papers in the list (based on the Geoscience Education Research Strength of Evidence Pyramid).

the rubric and can be used by atmospheric scientists as a resource for their teaching or as models for how to design and publish atmospheric science education research. We encourage readers to further explore these papers, and below we describe one representative paper from each level.

The papers that represent Practitioner Wisdom and Expert Opinion contribute pedagogical content knowledge to the research process, and "inform research by highlighting potential challenges, promising practices, and puzzling questions to address" (St. John and McNeal 2017, p. 367). At this level, authors are not explicitly conducting education research or analyzing results and may only make anecdotal observations and unsupported claims. As an example of a Level 1 paper, Coleman and Mitchell (2014) sought to increase student engagement by involving students in real-time data collection and analysis through high-altitude ballooning research. Although no direct quantitative evidence for improvement in student achievement was collected, the authors anecdotally observed that student attitudes and engagement improved and that program enrollments had increased. Papers in this category may have limited evidence to support claims of student learning, but they remain a rich source of evidence of the kind of teaching and instructional activities being promoted in the community and can lead to further research on novel and widespread education practices.

Papers in Level 2A represent quantitative and qualitative case studies that intentionally measure the effectiveness of particular pedagogies or instructional activities. At this level, researchers often use student self-reported data, such as project assignments, unit assessments, or student opinions from end-of-course evaluations. Therefore, these papers provide evidence representing the judgements and opinions of the participants. In Neves et al. (2013), the authors used an interactive modeling tool to model blackbody radiation and gradient wind dynamics for students in an introductory meteorology course. Using a Likert scale survey at the end of the course, they found that students considered the computational modeling activities useful for learning in meteorology and for their professional training as a whole. When systematically analyzed, data of this type can provide a general measure of learning gains, even if those gains may not be directly attributable to the pedagogy or learning activity being studied.

Providing evidence that demonstrates correlation between a pedagogy or innovation and its intended effect is a hallmark of Level 2B papers (and higher). Research at this level includes the intentional collection of both baseline and summative assessment data, often done through the administration of pre- and posttests or surveys. These data may be either quantitative, such as scored quizzes or concept inventories, or qualitative, including student quotes or reflections. For example, to assess the effectiveness of a field experience on student learning, Barrett and Woods (2012) collected and analyzed two years of pre- and postactivity questionnaires and content quizzes. Over both years, students demonstrated statistically significant learning gains in almost every content area studied. By using pretests to establish students' baseline knowledge, the authors were able to establish a correlation between their intervention and the measured gains. Demonstrating this correlation builds confidence in the ability of an instructional approach to improve the learning experience of students in a course. Authors may also use other methods to provide evidence of change, such as program evaluations or comparison between assessment methods.

The papers in Level 2C compare results to control groups or control scenarios, report on effect sizes, or use valid qualitative methods to establish the veracity of findings. These methods can provide evidence that an education practice was the cause of measured learning gains. Because humans are natural learners, time simply spent in a classroom along with student maturation can result in learning gains (Hattie 2012). Thus, in order to demonstrate the efficacy of education interventions with stronger evidence than gains demonstrated through pre- and posttests, it is essential to employ methods that establish causation.

One paper at this level, Grundstein et al. (2011), demonstrated causality of their intervention through a control scenario. To test the efficacy of a severe weather forecast activity, they divided students into traditional and inquiry-based laboratory sections. The control group completed exercises from a commonly used laboratory manual, while the experimental group performed a simulation in which they played the role of an operational forecaster on a severe weather day. Postlaboratory assessments showed that students in both groups demonstrated about the same level of content knowledge; however, students in the inquiry-based laboratory group demonstrated greater perceived learning and greater enjoyment of the laboratory exercise. The authors also reported representative written comments from the students, to support and clarify their statistical findings.

Quantitative and Qualitative Cohort Studies (Level 3) investigate a broad cross section of courses, institutions, and/or populations, and the findings resulting from these studies are supported with larger bodies of evidence collected from wide samples, making their claims more generalizable to overall atmospheric science education. The instruments used in these studies are broadly applicable such that the study can be easily replicated and the research is often conducted by external researchers, rather than course instructors, to minimize bias. Mackin et al. (2012) reported on a repertoire of rotating-tank experiments and curriculum in fluid dynamics involving 12 instructors and over 900 students, spread among 6 universities. The authors collected and analyzed instructor logs and surveys, held meetings with collaborators and faculty, and conducted site visits to observe the faculty in action. A validated pre- and posttest instrument was also developed (over 3 years) to assess students' content gains. The results demonstrated that the experiments fostered student curiosity, engagement, and motivation, expanded student cognitive models, and that in all cases experimental groups consistently showed greater learning gains on pre-/posttests than comparison groups.

We did not observe papers at Level 4 (Meta-Analyses) or Level 5 (Systematic Reviews). Completion of studies at these levels is dependent on a broad literature base within the foundational levels described above. As the community continues to address the need for increased participation in SoTL and for improved access to relevant literature on teaching and learning (Charlevoix 2008), we anticipate that additional studies at this level will emerge and provide greater insight into teaching and learning in undergraduate atmospheric science courses.

Discussion of findings

Our research question, "What is the current state of atmospheric science education research?" was designed to connect atmospheric science educators with existing literature and make transparent the strengths and limitations of the research that inform this body of work. Some may see ways to participate in the growing atmospheric science education research community, learning how to effectively evaluate and share their classroom innovations, while anyone may enrich their teaching with evidence-based practices. We know that many of our colleagues are innovating in the classroom and recognize that they may also wish to contribute to the published literature. We hope that providing names and descriptions of atmospheric science education research is helpful in planning those contributions.

A majority (25, or 53%) of the papers we reviewed were either Level 1 or Level 2A, and we anticipate additional contributions at this level as the atmospheric science education research community matures. Few papers (8, or 17%) were published at Level 2C or Level, which indicates there is also room for growth at these levels. The lack of papers at higher levels of the pyramid is not surprising or unsettling. St. John and McNeal (2017, p. 368) wrote that meta-analyses and systematic reviews are "the least common, in part, because they depend on access to data, methods, and findings from previously published research." To make investigations at Levels 4 and 5 possible, a sufficient body of high-quality work in Levels 1, 2, and 3 must be completed first.

We recognize that the authors of the majority of papers we reviewed are not trained in education research, lack the time and funding to invest in rigorous course or program evaluation, and do not have access to control groups and data from other programs. We found, however, through an examination of authors and their affiliations, that almost all papers nearing the upper echelons (Levels 2C and 3) had at least one coauthor with a background and/or formal training in education research. This is commensurate with patterns in other disciplines. Charlevoix (2008) offered a framework for investigators interested in carrying out their own SoTL projects but suggested SoTL research can be interdisciplinary, combining experts in atmospheric science and education research.

Because the topic of funding is central to strengthening atmospheric science education research, we identified funding sources, if provided by the author. Of the 47 papers included, less than half (47%) of the studies were at least partially funded. Eight papers were unclear about their funding source; for example, data in one study came from an NSF-sponsored Research Experience for Undergraduates (REU) but a statement of support or funding was not disclosed. Sources of funding ranged from governmental agencies such as NSF and NOAA, to internal funding from individual departments within universities. The most highly cited institution that provided funding was NSF (30%). This suggests that about half of the education research in atmospheric science is unsupported by formal funding sources. Significant funds are not necessarily required for SoTL studies, though costs of publications can be high for small, but valuable, studies.

We recognize that establishing a causal link between learning gains and an instructional intervention can be difficult, especially when all instruction is anticipated to result in learning gains (Hattie 2012). However, methods with the potential to establish causal links between interventions and learning can be intentionally chosen. Exam grades and end-of-course evaluations are not generally considered research data. They can provide some evidence but are limited in their ability to establish generalizability due to uncontrolled factors and, in the case of end-of-course evaluations, low response rates. Additionally, we found a wide range in how qualitative data were used. In multiple cases, authors used quotations from students or participants to illustrate points, but did not contextualize the representativeness of these quotations in comparison with other data. When reviewing these papers, we sometimes asked ourselves, “Were these comments cherry-picked?” Authors should be as transparent as possible when reporting qualitative data by contextualizing the circumstances in which the data were gathered and reporting how characteristic the selected examples are to the entire sample.

A limitation of our study is that we chose not to include conference presentations, abstracts, and preprints. By considering SoTL as scholarship, we focused on peer-reviewed sources. We acknowledge this choice led to the exclusion of a large body of work that is valuable to the atmospheric science education research community and that contributes to forming the base of the pyramid. Additionally, through our literature search process, we attempted to identify as many papers that fit our criteria as possible. Inevitably (perhaps due to key terms assigned to a paper, etc.) we missed some. For example, during the review process we became aware that Schultz (2010) should have been included but was not. There are likely others. Thus, our recommendations were made without the benefit that review of these papers may have offered.

Recommendations

Based on our findings and the work that has been accomplished thus far, we offer four recommendations for advancing the field of atmospheric science education research.

Recommendation 1. To grow the research base and encourage vertical development within the pyramid, support from the atmospheric science community is needed, especially at administrative and professional community levels. When commissioners and boards of

professional societies and heads and chairs of departments encourage and incentivize the development of atmospheric science education research, individuals are better positioned to seek opportunities to build and expand their science education research skills. AMS could support atmospheric science education research by promoting efforts and publicly aligning with NAGT to enable connections between organizations. Community-building efforts and sponsored programs (e.g., NSF, NOAA) could contribute toward generalizability by encouraging collaborations that help atmospheric scientists interested in education research collect stronger evidence.

Recommendation 2. By organizing and expanding efforts, atmospheric science education researchers can secure grants to support studies. NSF programs present opportunities for funding both in the NSF Division of Atmospheric and Geospace Sciences (e.g., GEOPaths), as well as in the Division of Undergraduate Education [e.g., Improving Undergraduate STEM Education (IUSE) and Building Capacity for STEM Education Research (BCSER)]. We encourage atmospheric science educators to apply for funding in atmospheric science education research, both internally through individual institutions and from larger funding sources. With grants and funding come studies with stronger evidence, publications, advancement of atmospheric science education research, and greater recognition for the value of education research within atmospheric science.

Recommendation 3. Collaborating with education researchers or evaluation specialists on project design, structure, and methods can move investigations into the realm of collecting causal evidence and improving atmospheric science education research. Some universities house science education specialists in science departments, and these individuals make ideal research partners. Colleges of education also include science educators who could provide an interdisciplinary approach to an atmospheric science education study. Atmospheric scientists interested in publishing on their teaching innovations might also consider joining NAGT, a network of professionals committed to high-quality, scholarly research in geoscience education. NAGT offers professional development and facilitates research partnerships.

Recommendation 4. We encourage the use of valid and reliable methods of evaluation when possible. The Geoscience Education Researcher Toolbox (<https://nagt.org/nagt/geodresearch/toolbox/>) includes a collection of instruments and surveys with validated and reliable measures for use in education research. Similar to other research areas, software packages for data analysis are customary and available to support quantitative and qualitative research designs (e.g., SPSS, R, MAXQDA, TAMS Analyzer). Education research often uses qualitative methods, which may necessitate some preparation on the part of researchers venturing into education research for the first time. Training on education research methods, instrumentation, data analysis, and publication can be found through resources hosted online, as well as through webinars, workshops, and conferences hosted by professional organizations such as AMS and NAGT.

Conclusions

The atmospheric science education community has built a strong literature foundation that contains a wealth of practitioner knowledge along with multiple case studies which provide evidence of the effectiveness of numerous education practices and innovative pedagogies. It is time to build upon this foundation to encourage continued development of atmospheric science education research. For example, we can use the education innovations described by practitioners (Level 1) and evaluate them as case studies. We can replicate case studies completed at Levels 2A and 2B using control groups and measure effect sizes or use qualitative

methods for deep inquiry. Replication will not only confirm initial results but provide stronger evidence upon which we can base claims. We can repeat case studies conducted at Level 2C in multiple classes, courses, or institutions to evaluate the generalizability of initial results. Ultimately, by adding to the literature at Level 2C and Level 3, we will provide a large enough base for conducting meta-analyses and systematic reviews. An intentional community effort to contribute at all levels of the pyramid is essential work that advances our field. It is our hope that scientists and educators can now envision how they might build upon the foundation of—and use—atmospheric science education research to ensure that we are training future atmospheric scientists in the best manner possible.

Acknowledgments. We thank Hannah Scherer and Karen McNeal for guidance on this project. We are grateful for the critiques of Dr. David Schultz and two anonymous reviewers. Their feedback significantly improved the manuscript and helped make education research accessible to the *BAMS* readership. This work was supported in part by a grant from the Indiana University Center for Innovative Teaching and Learning, Towson University, and by the National Science Foundation under Grant AGS-1560419.

Data availability statement. No datasets were generated or analyzed during the current study. The papers that we characterized are available at locations cited in the reference section.

References

- Barrett, B. S., and J. E. Woods, 2012: Using the amazing atmosphere to foster student learning and interest in meteorology. *Bull. Amer. Meteor. Soc.*, **93**, 315–323, <https://doi.org/10.1175/BAMS-D-11-00020.1>.
- Billings, B., S. A. Cohn, R. J. Kubesh, and W. O. J. Brown, 2019: An educational deployment of the NCAR Mobile Integrated Sounding System. *Bull. Amer. Meteor. Soc.*, **100**, 589–604, <https://doi.org/10.1175/BAMS-D-17-0185.1>.
- Bitting, K. S., R. Teasdale, and K. Ryker, 2018: Applying the geoscience education research strength of evidence pyramid: Developing a rubric to characterize existing geoscience teaching assistant training schedule. *J. Geosci. Educ.*, **65**, 519–530, <https://doi.org/10.5408/16-228.1>.
- Bond, N. A., and C. F. Mass, 2009: Development of skill by students enrolled in a weather forecasting laboratory. *Wea. Forecasting*, **24**, 1141–1148, <https://doi.org/10.1175/2009WAF2222214.1>.
- Cervato, C., W. Gallus, P. Boysen, and M. Larsen, 2011: Dynamic weather forecaster: Results of the testing of a collaborative, on-line educational platform for weather forecasting. *Earth Sci. Inf.*, **4**, 181–189, <https://doi.org/10.1007/s12145-011-0087-2>.
- Charlevoix, D. J., 2008: Improving teaching and learning through classroom research. *Bull. Amer. Meteor. Soc.*, **89**, 1659–1664, <https://doi.org/10.1175/2008BAMS2162.1>.
- Charlton-Perez, A. J., 2013: Problem-based learning approaches in meteorology. *J. Geosci. Educ.*, **61**, 12–19, <https://doi.org/10.5408/11-281.1>.
- Clements, C. B., and A. J. Oliphant, 2014: The California State University Mobile Atmospheric Profiling System: A facility for research and education in boundary layer meteorology. *Bull. Amer. Meteor. Soc.*, **95**, 1713–1724, <https://doi.org/10.1175/BAMS-D-13-00179.1>.
- Cobbett, E. A., E. L. Blickensderfer, and J. Lanicci, 2014: Evaluating an education/training module to foster knowledge of cockpit weather technology. *Aviat. Space Environ. Med.*, **85**, 1019–1025, <https://doi.org/10.3357/ASEM.3770.2014>.
- Cohen, A. E., and Coauthors, 2018: Bridging operational meteorology and academia through experiential education: The Storm Prediction Center in the University of Oklahoma classroom. *Bull. Amer. Meteor. Soc.*, **99**, 269–279, <https://doi.org/10.1175/BAMS-D-16-0307.1>.
- Coleman, J. S. M., and M. Mitchell, 2014: Active learning in the atmospheric science classroom and beyond through high-altitude ballooning. *J. Coll. Sci. Teach.*, **44**, 26–30, https://doi.org/10.2505/4/jcst14_044_02_26.
- Collins, R. L., S. P. Warner, C. M. Martus, J. L. Moss, K. Leelasakulthum, and R. C. Winnan, 2016: Studying the weather and climate of Alaska across a network of observers. *Bull. Amer. Meteor. Soc.*, **97**, 2275–2286, <https://doi.org/10.1175/BAMS-D-15-00202.1>.
- Croft, P. J., and J. Ha, 2014: The undergraduate “consulting classroom”: Field, research, and practicum experiences. *Bull. Amer. Meteor. Soc.*, **95**, 1603–1612, <https://doi.org/10.1175/BAMS-D-13-00045.1>.
- Davenport, C. E., 2018: Evolution in student perceptions of a flipped classroom in a computer programming course. *J. Coll. Sci. Teach.*, **47**, 30–35, https://doi.org/10.2505/4/jcst18_047_04_30.
- Drossman, H., J. Benedict, E. McGrath-Spangler, L. Van Roekel, and K. Wells, 2011: Assessment of a constructivist-motivated mentoring program to enhance the teaching skills of atmospheric science graduate students. *J. Coll. Sci. Teach.*, **41**, 72–81.
- Godfrey, C. M., B. S. Barrett, and E. S. Godfrey, 2011: Severe weather field experience: An undergraduate field course on career enhancement and severe convective storms. *J. Geosci. Educ.*, **59**, 111–118, <https://doi.org/10.5408/1.3604823>.
- Grundstein, A., J. Durkee, J. Frye, T. Andersen, and J. Lieberman, 2011: A severe weather laboratory exercise for an introductory weather and climate class using active learning techniques. *J. Geosci. Educ.*, **59**, 22–30, <https://doi.org/10.5408/1.3543917>.
- Harris, S. E., and A. U. Gold, 2018: Learning molecular behaviour may improve student explanatory models of the greenhouse effect. *Environ. Educ. Res.*, **24**, 754–771, <https://doi.org/10.1080/13504622.2017.1280448>.
- Hattie, J., 2012: *Visible Learning for Teachers: Maximizing Impact on Learning*. Routledge, 269 pp.
- Henderson, C., and Coauthors, 2017: Towards the STEM DBER alliance: Why we need a discipline-based STEM education research community. *Int. J. STEM Educ.*, **4**, 14, <https://doi.org/10.1186/s40594-017-0076-1>.
- Hepworth, K., C. E. Ivey, C. Canon, and H. A. Holmes, 2019: Embedding online, design-focused data visualization instruction in an upper-division undergraduate atmospheric science course. *J. Geosci. Educ.*, **68**, 168–183, <https://doi.org/10.1080/10899995.2019.1656022>.
- Hopper, L. J., Jr., C. Schumacher, and J. P. Stachnik, 2013: Implementation and assessment of undergraduate experiences in SOAP: An atmospheric science research and education program. *J. Geosci. Educ.*, **61**, 415–427, <https://doi.org/10.5408/12-382.1>.
- Horel, J. D., D. Ziegenfuss, and K. D. Perry, 2013: Transforming an atmospheric science curriculum to meet students’ needs. *Bull. Amer. Meteor. Soc.*, **94**, 475–484, <https://doi.org/10.1175/BAMS-D-12-00115.1>.
- Illari, L., and Coauthors, 2009: Weather in a tank: Exploiting laboratory experiments in the teaching of meteorology, oceanography, and climate. *Bull. Amer. Meteor. Soc.*, **90**, 1619–1632, <https://doi.org/10.1175/2009BAMS2658.1>.
- Jeong, J. S., D. González-Gómez, F. Cañada-Cañada, A. Gallego-Picó, and J. C. Bravo, 2019: Effects of active learning methodologies on the students’ emotions, self-efficacy beliefs and learning outcomes in a science distance learning course. *J. Technol. Sci. Educ.*, **9**, 217–227, <https://doi.org/10.3926/jotse.530>.
- Kahl, J. D. W., 2008: Reflections on a large-lecture, introductory meteorology course. *Bull. Amer. Meteor. Soc.*, **89**, 1029–1034, <https://doi.org/10.1175/2008BAMS2473.1>.
- , 2017: Automatic, multiple assessment options in undergraduate meteorology education. *Assess. Eval. Higher Educ.*, **42**, 1319–1325, <https://doi.org/10.1080/02602938.2016.1249337>.
- , and J. G. Ceron, 2014: Faculty-led study abroad in atmospheric science education. *Bull. Amer. Meteor. Soc.*, **95**, 283–292, <https://doi.org/10.1175/BAMS-D-13-00051.1>.
- Kelsey, E., C.-M. Briede, K. O’Brien, T. Padham, M. Cann, L. Davis, and A. Carne, 2015: Blown away: Interns experience science, research, and life on top of Mount Washington. *Bull. Amer. Meteor. Soc.*, **96**, 1533–1543, <https://doi.org/10.1175/BAMS-D-13-00195.1>.
- Kern, B., G. Mettetal, M. Dixon, and R. K. Morgan, 2015: The role of SoTL in the academy: Upon the 25th anniversary of Boyer’s Scholarship reconsidered. *J. Scholarship Teach. Learn.*, **15**, 1–14, <https://doi.org/10.14434/josotl.v15i3.13623>.
- Kopacz, D. M., 2017: Exploring refraction: Creating a superior mirage. *In the Trenches*, Vol. 7, No. 1, 9, <https://nagt.org/nagt/publications/trenches/v7-n1/v7n1p10.html>.
- , L. C. Maudlin, W. J. Flynn, Z. J. Handlos, A. Hirsch, and S. Gill, 2021: Involvement in and perception of atmospheric science education research. *Bull. Amer. Meteor. Soc.*, **102**, E717–E729, <https://doi.org/10.1175/BAMS-D-19-0230.1>.
- Kovacs, T., 2017: Experiencing the scientific method in a general education weather course. *In the Trenches*, Vol. 7, No. 1, 5, <https://nagt.org/nagt/publications/trenches/v7-n1/v7n1p5.html>.
- LaDue, N., P. McNeal, K. Ryker, K. St. John, and K. van der Hoeven Kraft, 2022: Using an engagement lens to model active learning in the geosciences. *J. Geosci. Educ.*, <https://doi.org/10.1080/10899995.2021.1913715>, in press.
- Laird, N. F., and N. D. Metz, 2020: A pair-researching approach for undergraduate atmospheric science researchers. *Bull. Amer. Meteor. Soc.*, **101**, E357–E363, <https://doi.org/10.1175/BAMS-D-18-0190.1>.
- Lanicci, J. M., 2012: Using a business process model as a central organizing construct for an undergraduate weather forecasting course. *Bull. Amer. Meteor. Soc.*, **93**, 697–709, <https://doi.org/10.1175/BAMS-D-11-00016.1>.
- Lerach, D. G., and N. R. Yestness, 2019: Virtual storm chase: Bringing the atmospheric sciences to life through an interactive case study. *J. Coll. Sci. Teach.*, **48**, 30–36, https://doi.org/10.2505/4/jcst19_048_03_30.

- Mackin, K. J., N. Cook-Smith, L. Illari, J. Marshall, and P. Sadler, 2012: The effectiveness of rotating tank experiments in teaching undergraduate courses in atmospheres, oceans, and climate sciences. *J. Geosci. Educ.*, **60**, 67–82, <https://doi.org/10.5408/10-194.1>.
- Mandrikas, A., D. Stavrou, and C. Skordoulis, 2017a: Teaching air pollution in an authentic context. *J. Sci. Educ. Technol.*, **26**, 238–251, <https://doi.org/10.1007/s10956-016-9675-8>.
- , —, and —, 2017b: A teaching-learning sequence about weather map reading. *Phys. Educ.*, **52**, 045007, <https://doi.org/10.1088/1361-6552/aa670f>.
- , —, K. Halkia, and C. Skordoulis, 2018: Preservice elementary teachers' study concerning wind on weather maps. *J. Sci. Teach. Educ.*, **29**, 65–82, <https://doi.org/10.1080/1046560X.2017.1423458>.
- Manduca, C. A., D. W. Mogk, and N. Stillings, 2003: Bringing research on learning to the geosciences: Report from a workshop sponsored by the National Science Foundation and the Johnson Foundation. Carleton College Science Education Resource Center, accessed 1 September 2020, https://serc.carleton.edu/research_on_learning/workshop02/.
- Market, P. S., 2006: The impact of writing area forecast discussions on student forecaster performance. *Wea. Forecasting*, **21**, 104–108, <https://doi.org/10.1175/WAF905.1>.
- Morss, R. E., and F. Zhang, 2008: Linking meteorological education to reality. *Bull. Amer. Meteor. Soc.*, **89**, 497–504, <https://doi.org/10.1175/BAMS-89-4-497>.
- Mullendore, G. L., and J. S. Tilley, 2014: Integration of undergraduate education and field campaigns: A case study from Deep Convective Clouds and Chemistry. *Bull. Amer. Meteor. Soc.*, **95**, 1595–1601, <https://doi.org/10.1175/BAMS-D-13-00209.1>.
- NAGT, 2020: Publishing SoTL vs DBER. Accessed 5 August 2020, https://nagt.org/nagt/geoedresearch/toolbox/publishing/sotl_dber.html.
- National Research Council, 2012: *Discipline-Based Education Research: Understanding and Improving Learning in Undergraduate Science and Engineering*. National Academies Press, 282 pp.
- Nelson, E. L., T. S. L'Ecuyer, A. L. Igel, and S. C. van den Heever, 2019: An interactive online educational applet for multiple frequencies of radar observations. *Bull. Amer. Meteor. Soc.*, **100**, 747–752, <https://doi.org/10.1175/BAMS-D-18-0249.1>.
- Neves, R. G. M., M. C. Neves, and V. D. Teodoro, 2013: Modellus: Interactive computational modelling to improve teaching of physics in the geosciences. *Comput. Geosci.*, **56**, 119–126, <https://doi.org/10.1016/j.cageo.2013.03.010>.
- Quardokus, K., S. Lasher-Trapp, and E. M. Riggs, 2012: A successful introduction of authentic research early in an undergraduate atmospheric science program. *Bull. Amer. Meteor. Soc.*, **93**, 1641–1649, <https://doi.org/10.1175/BAMS-D-11-00061.1>.
- Roberts, D., E. Bradley, K. Roth, T. Eckmann, and C. Still, 2010: Linking physical geography education and research through the development of an environmental sensing network and project-based learning. *J. Geosci. Educ.*, **58**, 262–274, <https://doi.org/10.5408/1.3559887>.
- Roebber, P. J., 2005: Bridging the gap between theory and applications: An inquiry into atmospheric science teaching. *Bull. Amer. Meteor. Soc.*, **86**, 507–518, <https://doi.org/10.1175/BAMS-86-4-507>.
- Scherer, H., C. Callahan, D. McConnell, K. Ryker, and A. Egger, 2019: How to write a literature review article for JGE: Key strategies for a successful publication. *Geological Society of America Annual Meeting*, Phoenix, AZ, GSA, Paper 264-4, <https://doi.org/10.1130/abs/2019AM-333357>.
- Schultz, D., 2010: A university laboratory course to improve scientific communication skills. *Bull. Amer. Meteor. Soc.*, **91**, 1259–1266, <https://doi.org/10.1175/2010BAMS3037.1>.
- , S. Anderson, and R. Seo-Zindy, 2013: Engaging Earth- and environmental-science undergraduates through weather discussions and an eLearning weather forecasting contest. *J. Sci. Educ. Technol.*, **22**, 278–286, <https://doi.org/10.1007/s10956-012-9392-x>.
- Shapiro, A., P. M. Klein, S. C. Arms, D. Bodine, and M. Carney, 2009: The Lake Thunderbird Micronet project. *Bull. Amer. Meteor. Soc.*, **90**, 811–824, <https://doi.org/10.1175/2008BAMS2727.1>.
- Shellito, C., 2020: Student-constructed weather instruments facilitate scientific inquiry. *J. Coll. Sci. Teach.*, **49**, 10–15, https://doi.org/10.2505/4/jcst20_049_03_10.
- St. John, K., and K. McNeal, 2017: The strength of evidence pyramid: One approach for characterizing the strength of evidence of geoscience education research (GER) community claims. *J. Geosci. Educ.*, **65**, 363–372, <https://doi.org/10.5408/17-264.1>.
- Tanamachi, R., D. Dawson, and L. C. Parker, 2020: Students of Purdue Observing Tornadoic Thunderstorms for Research (SPOTTR): A severe storms field work course at Purdue University. *Bull. Amer. Meteor. Soc.*, **101**, E847–E868, <https://doi.org/10.1175/BAMS-D-19-0025.1>.
- Yow, D. M., 2014: Teaching introductory weather and climate using popular movies. *J. Geosci. Educ.*, **62**, 118–125, <https://doi.org/10.5408/13-014.1>.