Environmental Science and Policy 137 (2022) 271–279

Contents lists available at ScienceDirect

Environmental Science and Policy

journal homepage: www.elsevier.com/locate/envsci

Development of a survey instrument to assess individual and organizational use of climate adaptation science

Steph L. Courtney [a,*], Amanda A. Hyman [b], Karen S. McNeal [a], Lindsay C. Maudlin [c], Paul R. Armsworth [b]

[a] Auburn University, Auburn, AL, USA
[b] The University of Tennessee, Knoxville, TN, USA
[c] Iowa State University, Ames, IA, USA

ARTICLE INFO

Keywords:
Climate change
Climate adaptation science
Science evaluation
Survey development

ABSTRACT

Research that can improve the resilience of social and natural systems to climate change has become more common. Many climate adaptation science organizations and agencies now focus on actionable science, a model that aims to have greater impacts on policy and practice than traditionally produced and distributed science. However, evaluations of research projects are needed to examine and verify the impact of climate science on adaptation and society. Better understanding the types and mechanisms of impact will allow organizations to design, fund, and facilitate more useful climate adaptation science. Many existing actionable science evaluation approaches are qualitative in nature and take considerable time and effort for funders and administrators to implement. Quantitative methods could provide a valuable option for evaluation, specifically for making comparisons across many projects. Thus, we have designed a quantitative survey instrument for measuring the use of climate adaptation science. We designed the survey using best practices and iterative input from social scientists as well as climate adaptation scientists and practitioners. We then distributed the survey to a sample of users of climate adaptation science and analyzed those responses to further refine the survey. Quantitative and qualitative results show that use of climate adaptation science may be described as either individual use or organizational use, which contrasts with popular models of use in existing evaluation literature. The survey is made available for future efforts to evaluate and improve climate adaptation science and to advance efforts to measure different kinds of use.

1. Introduction

Reducing harm from climate change will require varied and transformational adaptation responses, and such responses require knowledge and action across social and ecological systems (Fedele et al., 2019; Owen, 2020). In the last few decades, the fields of applied environmental science, especially climate adaptation science, have attempted to meet those needs by producing science that is more useful for decision-making and policy. However, making science more useful and evaluating whether those efforts have been successful are difficult and "messy" tasks (Arnott and Lemos, 2021; Nutley et al., 2007). This has inspired active research in both evaluation and improvement of climate adaptation science (Fazey et al., 2014; Louder et al., 2021). Many evaluation efforts have revealed that collaborative research practices, including but not limited to co-production and transdisciplinary science (Evely et al., 2010; Walter et al., 2007), often result in improved outcomes of climate adaptation science (Dilling and Lemos, 2011; Edwards and Meagher, 2020; Meadow et al., 2015; Owen, 2020).

Collaborative research practices contrast dramatically with the traditional method of knowledge exchange in academia, popularly called the loading dock approach (Cash et al., 2006). In the loading dock approach, scientists complete their research in isolation from the end-user and, when they are finished, deliver publications, websites, or (at best) decision-making tools to stakeholders. In this model, researchers seek to inform their stakeholders, however the created knowledge is less often used because it may not be at the spatial or temporal scale the manager needs, may not concern a priority topic, or may not be in a useful format or product, etc. (Bamzai-Dodson et al.,

* Corresponding author.
E-mail addresses: stephcourtneyphd@gmail.com (S.L. Courtney), ahyman2@vols.utk.edu (A.A. Hyman), ksm0041@auburn.edu (K.S. McNeal), maudlin@iastate.edu (L.C. Maudlin), p.armsworth@utk.edu (P.R. Armsworth).

https://doi.org/10.1016/j.envsci.2022.08.023
Received 12 March 2022; Received in revised form 23 August 2022; Accepted 28 August 2022
Available online 15 September 2022
1462-9011/

2021; Beier et al., 2017). Instead, co-production is one of many models that facilitates and encourages knowledge exchange, where "both researchers and stakeholders are now seen to have knowledge that is shared…" (pg. 2, Edwards and Meagher, 2020). However, different models and degrees of stakeholder involvement may be appropriate for different research efforts depending on the goals and constraints involved, especially since such involvement may require significantly more time, effort, and resources (Bamzai-Dodson et al., 2021; Meadow et al., 2015).

Understanding how research practices may lead to favorable results requires reliable methods for evaluating and measuring those results. There are a variety of outputs, outcomes, and impacts of climate adaptation science that can be measured, with some authors describing up to 16 types of outcomes and impacts (Wall et al., 2017). Usage of each term varies widely in the literature (Louder et al., 2021), but as commonly used in program evaluation, outputs include direct products delivered by the program; outcomes are changes in individuals' knowledge, skills, and behaviors; and impacts are larger changes in systems (Kellogg, 2004). Previous efforts to evaluate climate science projects have employed qualitative coding structures (Edwards and Meagher, 2020; Koontz et al., 2020), surveys (Walter et al., 2007), and citation metrics (Evely et al., 2010). Others have used literature reviews and interviews to build frameworks for measuring the use of collaborative climate science by resource managers and stakeholders (Edwards and Meagher, 2020; VanderMolen et al., 2020; Wall et al., 2017). In just the last few years, multiple authors have even meta-analyzed climate science evaluation literature to learn about research characteristics most often studied (Karcher et al., 2021; Koontz et al., 2020) and the evaluation frameworks applied (Louder et al., 2021; Reed et al., 2021).

One finding of those studies is that the usability of research may be the most frequent goal and evaluation metric used in climate adaptation science (Karcher et al., 2021). A common definition for usable science is that of Cash et al. (2003), wherein research that is credible, legitimate, and salient is more usable. However, Karcher et al. (2021) found that actual use of the science, rather than the theoretical usability of it, is less often mentioned as an explicit goal, claimed outcome, or topic of evaluation. This could be because the construct of use of science can be notoriously difficult to measure in any field, including climate adaptation science (Arnott and Lemos, 2021; Dilling and Lemos, 2011; Schwandt, 2015). In this study, we add to the literature by developing and testing a new tool for measuring the use of climate adaptation science.

1.1. Context and objectives

Many of the evaluation approaches mentioned above use qualitative methods which are useful for understanding research projects in depth. However, in some cases, it may be of interest for funders, policymakers, or organizations to compare many projects in less detail, for which there are few methods available in the literature. To address that absence, Hyman et al., 2022) conducted a study to develop a quantitative summative evaluation strategy for climate adaptation science projects and examine the relationships between project inputs, processes, and outcomes of 28 projects funded by the Southeast Climate Adaptation Science Center (SECASC). In that study, the authors sought to compare project characteristics that influenced two outcomes, scientific impact and use by partners. Scientific impact was measured via scientometrics including number of publications and citations. However, there was not a quantitative instrument available to measure stakeholder use of climate adaptation science. Thus, the present study was initiated to develop a survey to quantify use of climate adaptation science across many projects which was subsequently applied by Hyman et al., 2022 to understand the characteristics of climate adaptation science projects that may lead to greater use. We included use of any result or product of the project including publications, web tools, capacities built from engagement with the research process, workshops, etc.

The objectives of the current study were first, to develop a valid and reliable survey instrument to measure use of climate adaptation science, and second, to explore and examine any internal structure or sub-types of use that the instrument may reveal. We addressed these objectives by conducting a thorough and iterative survey development process, distributing the survey to a deliberate sample of partners, and analyzing the results of the survey.

This survey was not developed to measure any absolute degree of use of the projects. Instead, we sought to develop an instrument that could detect variation in use of research outcomes or products relative to other projects. In the future, the survey could be modified or used as-is for other research evaluation efforts, particularly by funding agencies that want to understand differences between large portfolios of projects using fewer institutional resources than qualitative methods would require (Bisbal, 2019). Better understanding the influences on completed projects can then inform future requests for proposals, funding decisions, and facilitation practices. Future applications of the survey could also be followed or accompanied by qualitative methods that would provide better understandings of why and how a project was used.

2. Methods

2.1. Survey Development

To address our first objective, we designed the survey following the systematic, seven-step process outlined by Artino et al. (2014), which includes several iterations of construct- and item-defining informed by literature review and interviews with the intended audience. Following an intentional process such as this one, which includes considerations from multiple perspectives and information sources, can provide greater evidence for the validity of a survey instrument (Artino et al., 2014; Libarkin et al., 2018). Throughout this process, the research team met monthly to discuss the developments and challenges at each step, and a smaller team met weekly for several months. Below, we summarize the seven steps of survey development and how we approached them (Table 1).

The first step of the process is to conduct a literature review, both to inform the survey development and to ensure there are not existing

Table 1
Survey development summary, adapted from Table 1 in Artino et al. (2014).

Survey design step	Summary of process and findings in this study
1. Conduct a literature review	Literature showed that no comprehensive quantitative instrument existed currently and that a new quantitative instrument was needed
2. Conduct group and individual interviews	Conducted interviews both at a scientific meeting and online via videocall afterwards, qualitative analysis and research group discussion were used to process and contextualize findings
3. Synthesize the literature review and interviews	Findings of step 2 roughly paralleled those of VanderMolen et al. (2020) so both were used to define indicators
4. Develop items	5-level Likert-style questions and examples were developed with input from the research team, data, and previous literature
5. Conduct expert validation	Items were reviewed, discussed, and refined by the research team and stakeholder advisory group
6. Conduct cognitive interviews	Respondents were interviewed while taking the survey to inform language and ordering adjustments
7. Conduct pilot testing	The 22-item survey was distributed to identified partners resulting in 81 complete responses
Additional follow-up interviews (not in Artino et al., 2014)	Interviews with survey respondents provided further insights into participant response interpretations, processes, and decisions; some results presented below

instruments that could be tested or adapted instead of creating a new instrument. Towards this end, we read relevant literature and compiled a list of items previously used in similar science evaluation efforts (Appendix A). Of the items that were explicitly provided in the literature, many addressed other constructs such as broad benefits, satisfaction, and partnerships. Of those that addressed use, usefulness, and effectiveness, the items were often open-ended or addressed the usability of one specific product. Thus, we decided there was a need for a new instrument to quantitatively measure the use of climate adaptation science projects, though the survey could likely be applied to or modified for other natural science uses.

The second step is to conduct interviews and/or focus groups to compare how the construct of interest is described and conceptualized by the potential survey respondents to how it is described in the scientific literature. To gather data for this step, we conducted group and individual interviews throughout a climate adaptation-focused scientific meeting. The interviews were separated by individuals' roles as either a science user, mostly including cultural and resource managers, or a science creator, including government and university researchers. For those who work in both capacities, we invited them to participate in multiple group interviews sharing from one perspective at a time. Because we were recruiting participants between and during meeting events, the interviews varied in length (from 10 to 40 min) and size (from two to six participants), with most group interviews lasting about 20 min with two to four participants. In total, we conducted eight group and two individual interviews with science producers and seven group and one individual interview with science users. Moving forward, we will use the term "partners" to describe the science user audience, acknowledging that individuals who use climate adaptation science hold a variety of professional and volunteer positions. At this stage, we asked participants how they have used climate adaptation science to reach their professional goals, what characteristics made the projects useful, and how their involvement and/or deliverables of the projects impacted their use (Courtney et al., in preparation). The recordings of the interviews and notes taken by the research team were transcribed and uploaded into qualitative analysis software (Dedoose). These data were inductively analyzed, primarily using content analysis, to identify major themes and concepts mentioned by participants and to examine the prevalence and connections between those themes (Krippendoff, 1989).

Given results in the first two steps, we chose to build upon the sub-constructs presented by VanderMolen et al. (2020) as a framework for our measurement of use. Specifically, those authors adapt the three most commonly described types of use (conceptual, instrumental, and justification) to stakeholder use of climate information. The coding criteria that VanderMolen et al. (2020) used to distinguish between types of use are: "information was reported to enhance knowledge base or to inform process or planning" (conceptual), "information was reported to influence decision-making directly with respect to action, process, or plan…" (instrumental), and "information was reported to justify an action, process, or plan…" (justification; p. 182, 2020), with examples corresponding to each. We developed operational definitions of each sub-construct which describe how the construct might be measured and are used to develop indicators for each construct, or a measurable occurrence that would indicate the presence or absence of the construct.

At this time, we conducted six more virtual individual interviews with natural and cultural resource managers involved with the SECASC to ensure our construct definitions were still compatible with our audience of interest. In this and following couple of stages, we did not intentionally include individuals in our final survey sample pool to avoid fatiguing those individuals. These interviews ranged from 20 to 40 min, and the questions focused on how climate science is used in participants' jobs, including questions centered on each of the three sub-constructs adapted from VanderMolen et al. (2020). These interviews were recorded, transcribed, and analyzed in the Dedoose software. The results of this stage were synthesized with previous literature review to develop indicators of each type of use, representing step three of the process. The indicators derived from our analysis paralleled the findings of VanderMolen et al. (2020).

For step four, the indicators were used to develop survey items (questions) to measure participants' degree of each use. Following the guidance of Artino et al. (2014) and previous survey experience, we chose to use quantity-based response options and varying operative verbs (e.g., influence, impact, been used to). For example, the item response options for questions about the project's influence on a particular decision were: 1. No influence, 2. Little influence, 3. Some influence, 4. Quite a bit of influence, and 5. A great deal of influence. For each question, we also compiled a short list of examples, which were presented with the questions, drawn from our own experiences and previous interviews (Table B.1).

The fifth step of survey development is to subject the preliminary questions to expert review. Toward this end, the questions were shared with our wider research team and stakeholder advisory group, whose members are experts in climate adaptation science, natural resource management, and climate science evaluation. After these changes were implemented, the survey was uploaded to the platform Qualtrics to confirm the visual design and for final review by the broader research team (Figure B.1).

Going forward, the survey items will be referenced by codes that begin with a letter where the letter refers to the theoretical construct the item was intended to represent. Thus, items that begin with a *C* were intended to represent conceptual use, *I* for instrumental use, *J* for justification use, and a number referring to their sequential ordering in the block. As the items were discussed in-depth by the research team, some that we initially assigned only to the conceptual use sub-construct became boundary items (i.e., we were no longer confident which sub-construct they would match best). Those items included one concerning monitoring efforts and two concerning education efforts, because each involves acting (instrumental) to gather or spread information (conceptual). Thus, we decided they were relevant to both conceptual and instrumental uses and created the *CI* combined sub-construct. Additionally, after the expert review stage, we added another response option to all but the first four items (C1-C4) that read *not applicable – my organization doesn't do this or I don't know*. Future users of this survey may want to change this language or separate the causes for selecting this response to receive more helpful information for their context.

Next, we recruited three individuals to interview while they took the survey for the sixth step. The goal of this step is to examine how the respondents "interpret the survey items and if their interpretation matches what the survey designer has in mind" (p. 470) to check for evidence of response process validity (Artino et al., 2014). We changed some of the question language and formatting of the items based on participant responses during this stage. The participants also noted that some questions were difficult to understand, specifically item J1 ("affirmed what you already know about environmental change and your job"). However, we retained the question because it is best survey practice to include imperfect items while testing a survey to gather more evidence for the deletion or inclusion of each.

2.2. Survey Distribution

We then distributed the survey to possible users of one or more of the 28 research projects of interest, fulfilling the seventh step of the survey design process (pilot distribution). The research projects examined were all funded beginning in the fiscal years from 2011 to 2016 with project durations ranging from one to five years. Each project was funded by the SECASC based on proposals from the research teams, who consisted primarily of university and federal scientists. We identified possible users of these research projects through project reports, publications, workshop attendance records, SECASC staff, the principal investigators of each project, and individuals named by other survey participants (snowball sampling). In the survey of principal investigators, we asked for the contact information of partners who would be able to comment on the use or usefulness of the research project and for the products, terms, or descriptors that partners would most likely associate with the project. Because we wanted to understand use of the projects relevant to decision-making, we excluded named partners who held research positions at universities. These partners held a variety of positions in settings including universities, non-governmental organizations, local, state, and tribal governments, and federal agencies such as the U.S. Fish and Wildlife Service. Some of the projects of interest involved stakeholders during research process (i.e., collaborative practices) and some did not; similarly, some of the stakeholder respondents were involved during the projects and many were not. This was a source of variation that was included in Hyman et al.'s analysis (2022).

In total, we identified 234 possible respondents and invited them to take the survey via email. In the survey invitation emails, we included a reference to the project which we believed them to be familiar with along with a webpage with the details of each project in case they were unaware of the formal titles or principal investigators associated with any project information they had used. The first page of the survey then prompted participants to select the project they were answering about, and on each page of the survey, instructions were included that read, "Please answer each question about the research project you selected and how you or the organization/agency you work for may have used information from it" (see Appendix B for survey visual design and complete instructions).

We sent two brief follow-up messages, spaced 1–2 weeks apart, to individuals who had not yet begun the survey by using the personalized link option. We received 81 complete responses (35%) which were used for the analyses described below.

We also invited participants to provide their contact information at the end of the survey for follow-up interviews to further understand and improve the instrument. In total, six participants provided their emails and were contacted following the survey completion. Only three participants responded to our follow up request and were interviewed to discuss their response process and understanding of the questions.

2.3. Analysis

Though *not applicable* responses will be meaningful and useful to future applications of the survey, they are not compatible with factor analyses. Thus, the *not applicable* responses in these data were deleted, treated as missing values, and imputed. The data were treated as ordinal for all analyses and imputed in MPlus software with the default settings for ordinal items (multiple imputation, polychoric correlations, and robust weighted least squares estimation; Asparouhov and Muthén, 2010; Jia and Wu, 2019; Muthén and Muthén, 1998–, 2017). The use of polychoric correlations can reduce the chances of over-dimensionalizing ordinal survey data in factor analysis (Van Der Eijk and Rose, 2015). This was important because the survey is meant to measure one overarching construct (use) with possible sub-constructs.

The survey was designed with a theoretical underlying structure in mind (conceptual, instrumental, and justification uses, as described above), meaning confirmatory factor analysis could have been used to test the fit of that structure. However, because this survey is one of the first attempts to quantify this structure, we instead decided to use exploratory factor analysis to address our second objective. Toward that end, we used parallel analysis with polychoric correlations (*random.polychor.pa* package in *R*; Presaghi and Desimoni, 2020) to determine the appropriate number of factors.

Exploratory factor analysis was conducted using the *psych* package in *R* version 4.1.1 (R Core Team, 2021; Revelle, 2021; RStudio Team, 2021). Once the appropriate number of factors was determined, we used oblimin rotation (a type of oblique rotation) to conduct the factor analysis. Because the survey was meant to uncover relationships within the overarching construct of use, any existing sub-factors should be correlated, making factor analysis and oblique rotation most appropriate (as opposed to principal components methods or orthogonal rotation). The factor analysis process was conducted iteratively, removing one item at a time, to simplify the structure (Watson, 2017). Specifically, items retained had one pattern coefficient (loading) at or above 0.60 and all others below 0.30 (the 0.6/0.3 rule; Matsunaga, 2010). Thus, retained items had strong loadings on only one factor and could be assigned to that factor with confidence. Although many authors use 0.40 as a minimum threshold for item retention, we decided it was appropriate to use stricter cut-offs because of our small sample size (Fabrigar and Wegener, 2012; Knetka et al., 2019; Watson, 2017). Once the final factor model was found, the *R* packages *lavaan* and *semTools* were used to calculate multiple measures of reliability for each scale (Jorgensen et al., 2021; Rosseel, 2012).

Though most of the qualitative data were used to generate and refine survey items and examples, some excerpts and results will be presented to contextualize our quantitative results. Additionally, the qualitative data were critical to improving the content validity and response process validity of the instrument while the quantitative analyses provide evidence of reliability (Artino et al., 2014; Libarkin et al., 2018).

3. Results

3.1. Survey results and factor analysis

The survey item median responses ranged from 1 to 4 (*little use* to *quite a bit of use*). Specifically, twelve items had a median of 3, six items with a median of 2, three items with a median of 1, and one item with a median of 4, each spread across constructs. All of the items were positively correlated with every other item, something to be expected given that we opted not to use any reverse response options, following recommended survey design practice (Artino et al., 2014). The polychoric (comparable to Pearson) correlation values ranged from 0.304 to 0.875 in the raw data and from 0.224 to 0.852 in the imputed data. The Kaiser-Meyer-Olkin factoring adequacy (KMO test) values ranged from 0.81 to 0.91 for each item with an overall value of 0.87, and Bartlett's test of sphericity was significant, which all indicate that the data were suitable for factor analysis (Watson, 2017).

In preliminary analyses, the parallel analysis showed that three factors best explained the variance in the dataset. However, the third factor was driven solely by J5 (support from lawmakers), the item with the second highest *not applicable* responses (30%). Thus, after more data exploration showing the influence of the imputed data, those items with $>$ 25% imputed data (*not applicable* responses) were removed from the dataset, which included items I6, J5, and I2 (Table 2). After these items were removed, the parallel analysis showed that only two factors were needed to explain the variance of the survey responses (initial model in Table 2). Exploratory factor analysis was run repeatedly in *R* using the.6/.3 rule, as described above, removing one item at a time. Once the items retained all satisfied the.6/.3 rule, analyses revealed that the data were still suitable for factor analysis and that two factors best explained the variance, so this model was retained as the final model (Table 3). The factors were named (individual and organizational use) based on the items that loaded onto them and findings from the qualitative data, discussed below.

Factor 1 (individual use) explained 33% of the data variance, and factor 2 (organizational use) explained 30%, resulting in a total R^2 of 0.63. The correlation between factors was 0.53. The Tucker-Lewis index of factoring reliability was 0.874; the model $\chi^2 = 75.7$ with $p < 0.001$; and the root mean square error of approximation was 0.122 (95% CI: 0.078–0.168). These values all fall outside of often-used thresholds for acceptable fit, which is a limitation of the model; however, improving global fit is not the goal of exploratory factor analysis, and the use of strict fit cut-offs has been criticized in recent years (Fabrigar and Wegener, 2012; Kline, 2016). Reliability analyses revealed that Cronbach's α was 0.90 for both factors and McDonald's ω was 0.92 for factor 1 and 0.93 for factor 2 (Hayes and Coutts, 2020; Kline, 2016).

Table 2
Survey items by factor (retained items) or reason for removal.

	Item	Item text: To what degree has this project…
Factor 1 Individual Use	C1	…impacted any of your professional skills?
	C2	…impacted your knowledge relevant to your job?
	C3	…influenced your professional network?
	C4	…changed your awareness of informational resources?
	C5	…influenced long-term planning documents?
	J2	…been used to encourage support from or collaboration with peers and/or partner organizations?
Factor 2 Organizational Use	C6	…influenced organizational/departmental objectives or priorities?
	C7	…influenced broad-scale or general policy?
	I3	…influenced decisions to change how time or labor are spent in your organization/agency?
	I4	…influenced decisions to change how money is allocated in your organization/agency?
	I5	…influenced decisions to change internal/organizational policies or procedures?
Removed: No loadings larger than or equal to 0.60	I1	…influenced decisions to change any habitat or species management practices?
	J4	…been used to encourage support from your supervisors?
	J1	…affirmed what you already know about environmental change and your job?
	J6	…been used to support a funding request?
Removed: Both loadings between 0.30 and 0.60	CI3	…influenced monitoring or research efforts in your organization or department?
	CI2	…impacted education efforts focused on resource managers or local landholders?
	CI1	…impacted public education efforts in your organization?
	J3	…been used to encourage support or cooperation from the public or local landholders?
Removed: over 25% *not applicable* responses	J5	…been used to encourage support from lawmakers?
	I6	…influenced decisions to change external policies, regulations, or enforcement?
	I2	…influenced decisions to change how infrastructure is managed?

Table 3
Exploratory factor analysis results for initial and final (retained) models.

Item (topic)	Initial Model Factor Loadings			Final Model Factor Loadings		
	Factor 1	Factor 2	h^2	Factor 1	Factor 2	h^2
C1 (skills)	**0.78**	-0.1	0.52	**0.74**	-0.02	0.53
C2 (knowledge)	**0.76**	-0.02	0.55	**0.76**	0.04	0.61
C3 (network)	**0.91**	-0.18	0.66	**0.85**	-0.09	0.64
C4 (information)	**0.82**	-0.16	0.54	**0.8**	-0.08	0.58
C5 (planning)	**0.68**	0.09	0.55	**0.74**	0.12	0.65
C6 (objectives)	0.24	**0.61**	0.61	0.25	**0.63**	0.63
C7 (policy)	0.18	**0.68**	0.64	0.16	**0.69**	0.62
CI1 (public ed.)	0.57	0.29	0.62			
CI2 (peer ed.)	0.56	0.34	0.66			
CI3 (monitoring)	0.36	0.48	0.58			
I1 (management)	0.48	0.19	0.38			
I3 (labor)	-0.04	**0.88**	0.73	-0.04	**0.86**	0.71
I4 (money)	0.05	**0.79**	0.67	0.04	**0.81**	0.69
I5 (procedure)	-0.08	**0.88**	0.7	-0.09	**0.88**	0.69
J1 (affirmation)	0.56	0.02	0.32			
J2 (partners)	**0.85**	-0.02	0.70	**0.62**	0.11	0.42
J3 (public)	0.59	0.32	0.67			
J4 (supervisor)	0.49	0.22	0.41			
J6 (funding)	0.53	0.19	0.45			

Note: Both models exclude items with > 25% not applicable responses. Bold denotes pattern coefficients > 0.60, h^2 is the communality of each item, and the empty boxes correspond to those items removed from the model.

3.2. Context from qualitative data

Each round of interviews we conducted had distinct prompts, purposes, and referenced different versions of the survey. Thus, while earlier rounds of interview were used to develop these survey questions, here we present context and quotes only from the last round of interviews, which referenced the survey in the form it was distributed, providing clarity in interpretation. While there were only three participants for this stage of interviews, the conversations with each ranged from 40 to 60 min, resulting in helpful information about participants' interpretations of and experiences with the survey.

One of the primary goals of the last round of interviews was to identify any difficulties participants had in answering the questions. The difficulties most often identified by the participants concerned language used in the questions. These language difficulties most often arose from mismatches between the questions and the reality of individuals' positions and uses of climate science. For example, one individual held a full-time volunteer position, so any mention of their job, professional skills, etc. did not technically fit their position. The same participant also found the question concerning infrastructure, which was ultimately removed due to high rates of *not applicable* responses, too vague ("You know, are we talking about physical infrastructure, but you could think maybe financial infrastructure or organizational infrastructure."). In actuality, we intended to measure impacts in any of these categories, so several questions were intentionally broad. Language disconnects like these demonstrate the challenges of developing a survey that is understandable to respondents in a wide range of positions, contexts, and backgrounds, which was also recognized by the same participant ("I don't really think you can get too much more specific just because of the really wide-spread audience that you're going to be applying the survey to. This comes along with a lot of gray area"). Only one participant highlighted a question where the question language did not align with their duties, which they described by sharing: "It was interesting, 'been used to encourage support or collaboration', I think I did struggle with the way that was worded a little bit. Because I didn't have to use anything to encourage support or collaboration, it's a constant thing to be involved with these groups."

Last, the project that one participant was responding about was completed less than a year before they took this survey, which influenced their response context. For some questions this made responding more difficult for them ("I'm pretty confident that within my [unit] it will influence some allocations, so I would've been torn at guessing at some impact versus I don't know, because it didn't actually happen yet. So it was a recurring problem through all the questions"). Notably, though, other questions with different operative words (though all in the past tense) were easier for this participant:

> "…they were all easy to answer because of the stage of the project and the verb tense. If not for the verb tense issue it would've been difficult to answer the last two because of the limited sphere that I can see. I don't know if it has been used for J5, or there might be an org using it for J6 that I don't know about. The way I answered them all was no impact because of the timing of the project."

Though unclear because only one participant mentioned it, differences in verbs between blocks could reduce the validity of the instrument which would present a limitation. However, this participant only drew a contrast between the use of "been used to" in the justification questions and all other verbs, which is not a major concern for the final instrument, because only one justification item (J2) was retained. Of the items retained in the final instrument, there are commonalities in the terms used by assigned factor, but the commonalities are not so uniform as to imply they are the reason for the factor loadings.

In contrast, other difficulties associated with question language may indicate complications that are inherent to measuring use of climate adaptation science. For example, all three participants described difficulties in deciding how to consider use at various scales and hierarchies of their organizations and in the context of their own positions, described by two of them in the quotes below (note: parenthetical text within excerpts represents interviewer speech and bracketed text represents edits made to maintain participant anonymity):

> "I struggled with the term organization. That's probably a bigger struggle for me than most people in my agency. My paycheck comes from one place where my job is to coordinate a multi-organization partnership, federal and state and private. I answer to the [partnership] more than the agency that pays me. I'd like to say I answered consistently but I can't promise I did. I tried to answer from the [partnership] perspective because that's how I was related to this project."

> "It's interesting, any federal agency is a juggernaut. To change course or even influence internal policies or organizational policies takes a thermonuclear weapon sometimes. Climate change is impacting the [agency] and we're seeing those changes happening slowly. So it kind of depends on what we're talking about. Did this particular project affect any changes to the internal policies or internal organization of the agency, nah, probably not. But it's one of many that are contributing to change, recognition, and options, what we need out here."

To some degree, these quotes describe dynamics common to large organizations, i.e., operations at multiple scales where information needs and applications vary. Additional quotes from participants seemed to show that there were two primary factors making it difficult for them to answer the questions about each of these scales. The first was, understandably, a single individual not knowing what information is useful or used at other locations or hierarchical levels of their organization, as described by the participants below:

> "There are some of these things that you don't know for sure, we've talked and hit upon this, the influence decisions in a larger perspective kind of questions – you can guess at it, maybe you know if you're involved with some of the larger groups in climate change or you get a call from the director in D.C. so you have an idea then, but otherwise you just don't have the foundation to answer about the larger organization."

> "This would be true of any of those projects on the list, there are influences we can identify and influences that we don't know about. I have a sphere that I can see and can answer about, but there are lots of influences outside that sphere, so that makes it hard to answer these."

> "I don't know how anyone could answer those last two besides I don't know, because anyone could do those and you would'nt necessarily be aware. You could add 'funding request by my department' or 'that I know of'."

In designing the survey, of course we only expected participants to answer from their knowledge and perspective and would not expect them to know how information is used across the entirety of large organizations or federal agencies. These quotes illustrate that our expectation that participants respond from their perspective may be clearer, as the last quote suggests, by explicitly asking about use "that they know of."

In addition to not having full awareness of how information is used at different scales of an organization, participants also described variation in the importance of one piece of information across scales. This variation complicated the response process for at least one of the participants, largely because the survey asked participants to try to quantify use via the Likert-style response options, as described below:

> "If I had to re-answer now, I'd say the study in itself probably did not have a tremendous impact, probably would've said little or some. Here at the [site], lots. So when it's broad like that it's harder to answer… I would say, if this had said "locally," at the [site] level, I would've without fail said that there was a great deal of impact… I think it's important to look at where the impacts occurred, especially if you're looking at how the study impacted the broader spectrum of things, or any study – for local managers and people with local knowledge, something like this can be huge. On a Washington-level scale, this is just one piece of many studies and many pieces of work that have gone into painting the whole climate change picture."

Participant responses also seem to indicate meaningful differences in how climate science is used across organizational scales, beyond the differences in awareness and importance described above. Two of the participants explicitly described this contrast by both separating use by an individual and organization and by comparing the conceptual items that were eventually assigned to the first factor (C1-C5) versus the second factor (C6-C7):

> "So when I answered J1, I would've been thinking about the whole process and not just outcomes. (Why J1?) Similar to the first three or four questions, how they impacted me, J1 is also a very personal questions, about me and my job. So that's why I thought about how the project impacted me, which was throughout the course of the whole project. It didn't impact my organization throughout the whole project except through me and my skillsets."

> "…like C5 – when it's completed I know it absolutely will influence those documents. I don't need it to be finished to know how influential it will be in our planning documents. It's harder to say how it will play out for C6. (Why?) Because I get to decide what goes in planning documents and a whole bunch of other people get to decide what impact they have. (So is it about your position?) Yes. And that's even more true with C7 – policy for my organization is set at a very high level and not by me or my boss."

> "C's seem more personal now that I compare them, where the second chunk are more similar to the last few questions of the previous chunk. The early C's are personal, then moved to organizational influence, and these all seem like organizational influence questions."

Participants also mentioned possible explanations of such differences across levels of an organization such as differences in funding structures and the breadth of factors under consideration ("I don't know that there's any level above that people are thinking about climate change… So that's probably the broadest that we're going to get"). When asked if the survey left out any ways that they use climate science, only one respondent had an answer which was use for media relations including applications to broadcast, print, and social media, a topic that could be added to the survey in the future.

4. Discussion

4.1. Evidence of instrument validity and reliability

We followed a systematic development process to ensure we had multiple opportunities to examine and improve the validity and reliability of the survey instrument. There are multiple definitions and types of validity described in the literature, but here we are using the definition provided by the *Standards for Educational and Psychological Testing*: "Validity refers to the degree of which evidence and theory support the interpretations of the test score for the proposed use" (AERA, APA, and NCME, 2014, p.11). The content validity was bolstered by the use of multiple rounds of interviews with the intended audience and including input from previous literature and content experts (Knekta et al., 2019; Libarkin et al., 2018). Because instrument validity is also reliant on the audience and context of use, repeatedly checking in with our intended users, natural and cultural resource managers engaged with the Southeast Climate Adaptation Science Center, was also important.

The last two sets of interviews which used preliminary and final versions of the survey questions provided evidence of response process validity by assessing how respondents interpreted and answered the questions (Artino et al., 2014). Exploratory factor analysis was conducted to evaluate construct dimensionality and validity based on internal structure, though the factor structure did not align with the initially intended sub-types of use. However, multiple participants in two of the last three rounds of interviews, where we were asking about ways that they use climate adaptation science, described differences in use that were very compatible with the resulting factor structure. We used conservative factor analysis methods (e.g., >0.6 factor loadings) to ensure that there was solid evidence for the survey structure despite our middling sample size, and the resulting factors had high reliability coefficients.

The results of Hyman et al. (2022) provide evidence of the validity of the survey via relationships to other variables. Those authors used the same survey responses as this study but ran independent analyses, resulting in a similar 2-factor structure. Those factors, their sub-types of use, were then used as outcomes in a structural equation model testing the relationships between various inputs to the projects and outcomes, the two types of use and academic impact. Their analysis found that the three outcomes did not share predictors and, instead, that the frequency of meetings between researchers and users significantly predicted one sub-type of use which then predicted the other use. Their analysis also found important relationships between other project characteristics (e.g., project budget and duration) and research publications. These findings, specifically the relationship between team meetings and use, align with previous research demonstrating the crucial role of consistent stakeholder engagement in collaborative science approaches for increasing use (Djentonin and Meadow, 2018). Additionally, the distinct influencing variables on the two sub-types of use suggest they are indeed separate and unique constructs.

4.2. Survey structure and implications

We developed the survey based on three sub-types of use (conceptual, instrumental, and justification; CIJ) which have been applied throughout previous literature in evaluation and often applied in climate adaptation science evaluation (e.g., Arnott and Lemos, 2021; Louder et al., 2021; Reed et al., 2014; in combination with other sub-constructs, also Edwards and Meagher, 2020). These authors all found additional evidence of the applicability of the CIJ sub-constructs, unlike the present study. There are many possible causes for the discrepancies in their findings and our expected findings (the CIJ structure) and our two-factor results. First, we note that each of the above examples have relied on qualitative analyses. It is therefore possible that organizing structures of use (i.e., CIJ versus individual and organizational) are related to the methods of investigation and evaluation. That said, these findings are not the first, qualitative or quantitative, to reveal differences in use of climate science based on institutional level or position of use. Cvitanovic et al. (2018) conducted a case study evaluation of an applied environmental science program and found distinctions between impacts on individuals, the host university, and policy and decision-making. Notably, some quotes the authors provide describing types of individual impacts (learning opportunities, expansion of social networks) and university/organizational impacts (relevance of research to policy) parallel those described by participants in this study. However, we do not disregard the rich previous literature illustrating the importance of describing conceptual, instrumental, and justification uses. Instead, our findings represent interesting evidence for broadening our understandings and measurement of use of climate adaptation science.

There are likely important connections between the CIJ and individual-organizational (IO) structures. For example, VanderMolen et al. (2020) posit that their findings may be a result of organizational cultural factors, specifically that conceptual use of information may predominate in government agencies. In the present study, we found that the survey items we intended to represent conceptual use largely loaded together, except for the last two items which some participants said represented different uses (less "personal"). Additionally, the participant descriptions of individual and organizational use, or use higher up the chain of command, highlighted some differences congruent with the CIJ structure. For example, because instrumental use involves changes in practices, institutional positionality would be very relevant to any individuals' capacity to change practices based on new research. In this study, participants noted that they considered setting priorities and policies (C6 and C7) to be broader tasks only carried out "at a very high level." Finally, the items that we placed in the boundary category, CI, had three of the five smallest differences between loadings on the two factors (cross-loadings) and thus were boundary items in the factor model as well.

Because this was one of the first efforts to quantify use of climate adaptation science, each step of the process was needed to build evidence for the validity of the survey. There are many organizations with similar research projects and partners, however, who may find it useful in its current form. Future implementations of the survey as presented could bolster the evidence of the reliability and validity of the instrument. Additionally, future studies could apply analyses such as a confirmatory factor analysis comparing the CIJ and IO models to enhance our collective understanding of how climate adaptation science is used and how to quantify that use. Alternately, the relationships between the models could be used to examine whether a combined or nested model might best explain use of climate adaptation science (i.e., conceptual organizational use vs. conceptual individual use). Regardless, understanding the finer details between types of use and when they emerge will likely require a much larger sample size than our 81 participant responses. Additionally, while we chose participants based on project documentation and researchers' recommendation, testing other sampling methods may provide different kinds of information about use of the projects, such as deliberate sampling of individuals at different managerial or authoritative levels of an organization.

Combining this survey with other evaluation methods and frameworks, for example those centered on stakeholder engagement (Bamzai-Dodson et al., 2021; Meadow et al., 2015) or on qualitative data which could provide more details about nuances and mechanics of use, could reveal how to increase use of climate adaptation science.

The development process we followed was helpful for building an instrument to fit our needs and respondents, but future implementations may benefit from making some adjustments to the survey based on the goals and context at hand. For example, practitioners could repeat some steps of the development process, especially step 6 (Table 1), to tailor the instrument to specific respondents. Second, items that were removed from the survey in this study could still be used in future implementations. In particular, items that were removed only for weak loadings (Table 2) or items that are highly relevant to the research being evaluated could be valuable additions. Lastly, based on our qualitative findings, it may be appropriate to change some of the language of the questions in the future. For example, evaluators could provide explicit definitions about the time and organizational perspectives that respondents should answer from (i.e., use by only you or including use by your coworkers, use in the last year, use that you are aware of, etc.) depending on the survey purpose and context. Of course, any of these modifications are also dependent on the time and labor available to dedicate to additional review, interviews, and/or data analysis to inform the changes.

This instrument and both models of use may be important for better understanding the impacts of climate adaptation science. Specifically, this quantitative survey can be useful for comparing larger suites of projects with less institutional time and funding than in-depth qualitative evaluations. Better understanding large suites of projects using faster evaluation methods may be important for improving climate adaptation science more broadly. To name a few examples, funding agencies and organizations can develop uniform and transferable evaluation protocols; alter requests for funding proposals (RFPs) to better represent their missions and priorities; find consistent project weaknesses to improve facilitation procedures; or tailor their research portfolios to be more targeted or more diversified (Arnott et al., 2020; Bisbal, 2019; Karcher et al., 2021).

5. Conclusions

We have presented a robust and thorough process for developing a survey which measures use of climate adaptation science. The results from the first distribution of the survey imply that our quantitative measures of use may not fit the previously theorized structures. Instead, the survey items seemed to covary depending on whether the item described use for individual purposes or at broader organizational scales. Interview data, collected throughout the survey development process, were used to discern and illustrate this difference. Enhanced understandings of how climate science influences policy and practice allow funding agencies to build more efficient and deliberate actionable science research portfolios. Ultimately, by understanding and improving the connections between science on society, we can more effectively and efficiently adapt to climate change to reduce harm to ecosystems and society.

Funding

The lead author was supported by the National Science Foundation, grants DGE-1414475 and DGE-1922687, and the study was funded by the United States Geological Survey, grant No. G20AC00121.

CRediT authorship contribution statement

S. Courtney: Data curation, Formal analysis, Investigation, Methodology, Resources, Software, Validation, Visualization, Writing – original draft, Writing – review & editing. **A. Hyman:** Investigation, Methodology, Software, Validation, Writing – review & editing. **K. McNeal**: Conceptualization, Funding acquisition, Methodology, Project administration, Resources, Supervision, Writing – review & editing. **L. Maudlin**: Investigation, Methodology, Validation, Writing – review & editing. **P. Armsworth:** Funding acquisition, Project administration, Resources, Supervision, Writing – review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data Availability

The data used in this study will be published on ScienceBase.

Acknowledgements

We would like to acknowledge Cari Furiness, Aranzazu Lascurain, Mitch Eaton, Aparna Bamzai-Dodson, Lalasia Bialic-Murphy, and our stakeholder advisors for their invaluable contributions and assistance throughout this study. We also appreciate our collaborators in the cross-CASC evaluation network for their feedback at key points in the project. This study was approved by the Auburn University Institutional Review Board (protocol #19–441 EX 1910).

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.envsci.2022.08.023.

References

American Educational Research Association, American Psychological Association, National Council on Measurement in Education, 2014. Standards for Educational and Psychological Testing. AERA.
Arnott, J.C., Kirchhoff, C.J., Meyer, R.M., Meadow, A.M., Bednarek, A.T., 2020. Sponsoring actionable science: what public science funders can do to advance sustainability and the social contract for science. Current Opinion in Environmental Sustainability 42, 38–44. https://doi.org/10.1016/j.cosust.2020.01.006.
Arnott, J.C., Lemos, M.C., 2021. Understanding knowledge use for sustainability. Environ. Sci. Policy 120, 222–230. https://doi.org/10.1016/j.envsci.2021.02.016.
Artino, A.R., La Rochelle, J.S., Dezee, K.J., Gehlbach, H., 2014. Developing questionnaires for educational research: AMEE guide no. 87. Med. Teach. 36 (6), 463–474. https://doi.org/10.3109/0142159X.2014.889814.
Asparouhov, T., & Muthén, B. (2010). Multiple imputation with Mplus. *MPlus Web Notes*. ⟨http://statmodel2.com/download/Imputations7.pdf⟩.
Bamzai-Dodson, A., Cravens, A.E., Wade, A., McPherson, R.A., 2021. Engaging with stakeholders to produce actionable science: a framework and guidance. Weather Clim. Soc. 13 (4), 1027–1041. https://doi.org/10.1175/wcas-d-21-0046.1.
Beier, P., Hansen, L.J., Helbrecht, L., Behar, D., 2017. A how-to guide for coproduction of actionable science. Conserv. Lett. 10 (3), 288–296. https://doi.org/10.1111/conl.12300.
Bisbal, G.A., 2019. Practical tips to establish an actionable science portfolio for climate adaptation. Sci. Public Policy 46 (1), 148–153. https://doi.org/10.1093/scipol/scy070.
Cash, D.W., Clark, W.C., Alcock, F., Dickson, N.M., Eckley, N., Guston, D.H., Jä Ger, J., Mitchell, N.M., 2003. Knowledge systems for sustainable development. Proc. Natl. Acad. Sci. 100 (14), 8086–8091. https://doi.org/10.1073/pnas.1231332100.
Cash, D.W., Borck, J.C., Patt, A.G., 2006. Countering the loading-dock approach to linking science and decision making: comparative analysis of El Niño/Southern Oscillation (ENSO) forecasting systems. Sci. Technol. Hum. Values 31 (4), 465–494. https://doi.org/10.1177/0162243906287547.
Courtney, S.L., Hyman, A.H., McNeal, K.S., Armsworth, P.R. (in preparation). Natural and cultural resource managers' use of climate science. Sciencebase.
Cvitanovic, C., Löf, M.F., Norström, A.V., Reed, M.S., 2018. Building university-based boundary organisations that facilitate impacts on environmental policy and practice. PLoS ONE 13 (9), 1–19. https://doi.org/10.1371/journal.pone.0203752.
Dilling, L., Lemos, M.C., 2011. Creating usable science: opportunities and constraints for climate knowledge use and their implications for science policy. Glob. Environ. Change 21 (2), 680–689. https://doi.org/10.1016/j.gloenvcha.2010.11.006.
Djentonin, I., Meadow, A., 2018. The art of co-production of knowledge in environmental sciences and management: lessons from international practice. Environ. Manag. 61 (6), 885–903. https://doi.org/10.1007/s00267-018-1028-3.
Edwards, D.M., Meagher, L.R., 2020. A framework to evaluate the impacts of research on policy and practice: a forestry pilot study. For. Policy Econ. 114 (May), 101975 https://doi.org/10.1016/j.forpol.2019.101975.
Evely, A.C., Fazey, I., Lambin, X., Lambert, E., Allen, S., Pinard, M., 2010. Defining and evaluating the impact of cross-disciplinary conservation research. Environ. Conserv. 37 (4), 442–450. https://doi.org/10.1017/S0376892910000792.
Fabrigar, L.R., Wegener, D.T., 2012. Exploratory Factor Analysis. Oxford University Press.
Fazey, I., Bunse, L., Msika, J., Pinke, M., Preedy, K., Evely, A.C., Lambert, E., Hastings, E., Morri, S., Reed, M.S., 2014. Evaluating knowledge exchange in interdisciplinary and multi-stakeholder research. Glob. Environ. Change 25 (1), 204–220.
Fedele, G., Donatti, C.I., Harvey, C.A., Hannah, L., Hole, D.G., 2019. Transformative adaptation to climate change for sustainable social-ecological systems. Environ. Sci. Policy 101 (May), 116–125. https://doi.org/10.1016/j.envsci.2019.07.001.
Hayes, A.F., Coutts, J.J., 2020. Use omega rather than Cronbach's alpha for estimating reliability. But…. Commun. Methods Meas. 14 (1), 1–24. https://doi.org/10.1080/19312458.2020.1718629.
Hyman, A.A., Courtney, S., McNeal, K.S., Bialic-Murphy, L., Furiness, C., Armsworth, P., 2022. Distinct pathways to stakeholder use versus academic contribution in climate adaptation research. Conserv. Lett. 15 (4), 1–10. https://doi.org/10.1111/conl.12892.
Jia, F., Wu, W., 2019. Evaluating methods for handling missing ordinal data in structural equation modeling. Behav. Res. Methods 51 (5), 2337–2355. https://doi.org/10.3758/s13428-018-1187-4.
W.K. Kellogg Foundation, 2004. *Logic Model Development Guide*.
Jorgensen, T.D., Pornprasertmanit, S., Schoemann, A.M., & Rosseel, Y. (2021). *semTools: Useful tools for structural equation modeling*. R package version 0.5–5. ⟨https://CRAN.R-project.org/package=semTools⟩.
Karcher, D.B., Cvitanovic, C., Colvin, R.M., van Putten, I.E., Reed, M.S., 2021. Is this what success looks like? mismatches between the aims, claims, and evidence used to demonstrate impact from knowledge exchange processes at the interface of environmental science and policy. Environ. Sci. Policy 125, 202–218. https://doi.org/10.1016/j.envsci.2021.08.012.
Kline, R.B., 2016. Principles and Practice of Structural Equation Modeling, fourth ed. The Guilford Press.

Knekta, E., Runyon, C., Eddy, S., 2019. One size doesn't fit all: Using factor analysis to gather validity evidence when using surveys in your research. CBE Life Sci. Educ. 18 (1), 1–17. https://doi.org/10.1187/cbe.18-04-0064.

Koontz, T.M., Jager, N.W., Newig, J., 2020. Assessing collaborative conservation: a case survey of output, outcome, and impact measures used in the empirical literature. Soc. Nat. Resour. 33 (4), 442–461. https://doi.org/10.1080/08941920.2019.1583397.

Krippendoff, K., 1989. Content analysis. In: Barnouw, E., Gerbner, G., Schramm, W., Worth, T.L., Gross, L. (Eds.), International Encyclopedia of Communication, Volume 1. Oxford University Press, New York, NY, pp. 403–407. ⟨http://repository.upenn.edu/asc_papers/226⟩.

Libarkin, J.C., Gold, A.U., Harris, S.E., Mcneal, K.S., Bowles, R.P., 2018. A new, valid measure of climate change understanding: associations with risk perception. Clim. Change 150, 403–416.

Louder, E., Wyborn, C., Cvitanovic, C., Bednarek, A.T., 2021. A synthesis of the frameworks available to guide evaluations of research impact at the interface of environmental science, policy and practice. Environ. Sci. Policy 116, 258–265. https://doi.org/10.1016/j.envsci.2020.12.006.

Matsunaga, M., 2010. How to factor-analyze your data right. Int. J. Psychol. Res. 3 (1), 97–110.

Meadow, A.M., Ferguson, D.B., Guido, Z., Horangic, A., Owen, G., Wall, T., 2015. Moving toward the deliberate coproduction of climate science knowledge. Weather Clim. Soc. 7 (2), 179–191. https://doi.org/10.1175/WCAS-D-14-00050.1.

Muthén, L.K. and Muthén, B.O. (1998–2017). *Mplus User's Guide*. Eighth Edition. Muthén & Muthén. ⟨https://www.statmodel.com/HTML_UG/introV8.htm⟩.

Nutley, S.M., Walter, I., Davies, H.T.O., 2007. How can we assess research use and wider research impact? Using Evidence: How Research Can Inform Public Services. Bristol University Press, pp. 271–296.

Owen, G., 2020. What makes climate change adaptation effective? a systematic review of the literature. Glob. Environ. Change 62 (April), 102071. https://doi.org/10.1016/j.gloenvcha.2020.102071.

Presaghi, F., & Desimoni, M. (2020). *A Parallel Analysis with Polychoric Correlation Matrices* (Version 1.1.4–04). ⟨https://cran.r-project.org/package=random.polychor.pa⟩.

R Core Team (2021). *R: A language and environment for statistical computing* (Version 4.1.1). R Foundation for Statistical Computing. ⟨https://www.R-project.org/⟩.

Reed, M.S., Ferré, M., Martin-Ortega, J., Blanche, R., Lawford-Rolfe, R., Dallimer, M., Holden, J., 2021. Evaluating impact from research: a methodological framework. Res. Policy 50 (4). https://doi.org/10.1016/j.respol.2020.104147.

Reed, M.S., Stringer, L.C., Fazey, I., Evely, A.C., Kruijsen, J.H.J., 2014. Five principles for the practice of knowledge exchange in environmental management. Journal of Environmental Management 146, 337–345. https://doi.org/10.1016/j.jenvman.2014.07.021.

Revelle, W., 2021. psych: Procedures for Psychological, Psychometric, and Personality Research (Version 2.1.9). Northwestern University. ⟨https://CRAN.R-project.org/package=psych⟩.

Rosseel, Y., 2012. lavaan: an R package for structural equation modeling. J. Stat. Softw. 48 (2), 1–36. https://doi.org/10.18637/jss.v048.i02.

RStudio Team, 2021. RStudio: Integrated Development Environment for R. RStudio. PBC. ⟨http://www.rstudio.com/⟩.

Schwandt, Thomas A., 2015. Evaluation foundations revisited: cultivating a life of the mind for practice. Stanford University Press.

Van Der Eijk, C., Rose, J., 2015. Risky business: Factor analysis of survey data - assessing the probability of incorrect dimensionalisation. PLoS ONE 10 (3), 1–32. https://doi.org/10.1371/journal.pone.0118900.

VanderMolen, K., Meadow, A.M., Horangic, A., Wall, T.U., 2020. Typologizing stakeholder information use to better understand the impacts of collaborative climate science. Environ. Manag. 65 (2), 178–189. https://doi.org/10.1007/s00267-019-01237-9.

Wall, T.U., Meadow, A.M., Horganic, A., 2017. Developing evaluation indicators to improve the process of coproducing usable climate science. Weather Clim. Soc. 9 (1), 95–107. https://doi.org/10.1175/WCAS-D-16-0008.1.

Walter, A.I., Helgenberger, S., Wiek, A., Scholz, R.W., 2007. Measuring societal effects of transdisciplinary research projects: design and application of an evaluation method. Eval. Program Plan. 30 (4), 325–338. https://doi.org/10.1016/j.evalprogplan.2007.08.002.

Watson, J.C., 2017. Establishing evidence for internal structure using exploratory factor analysis. Meas. Eval. Couns. Dev. 50 (4), 232–238. https://doi.org/10.1080/07481756.2017.1336931.

Steph L. Courtney, PhD, is the Assistant Climate Science Liaison with United South and Eastern Tribes Inc. (USET). She completed her PhD in Earth System Science at Auburn University with research exploring how individuals perceive and apply climate change science. To date, she has used eye-tracking, survey, interview, and statistical methods to better understand the impacts of climate change graph designs, perceptions of climate scientists, evaluation of climate adaptation science, and museum visitors' experiences using climate change information.

Amanda A. Hyman, PhD, is a Postdoctoral Researcher in the Department of Biological Sciences at the Virginia Polytechnic Institute and State University. She currently works with USGS's National Climate Adaptation Science Center to develop a multispecies indicator of species vulnerability to climate change. Her work is highly interdisciplinary spanning ecology, human dimensions, and international development, while studying anything from lions in Tanzania to African mahogany in Benin to rainbow trout in Virginia. Amanda addresses messy and interesting questions rather than focusing on specific systems; her work ultimately hopes to make conservation planning and funding efforts more efficient and effective.

Karen S. McNeal, PhD, is the Molette Endowed Professor in the Department of Geosciences at Auburn University in Auburn, Alabama. She is a discipline-based education researcher and studies how people think and learn about Earth systems and global change phenomena. Her work employs psychomotor tools (e.g., eye-tracking, pupillometry, and skin biosensors), psychometric tools (e.g., concept inventories and pre-post surveys), and open-ended approaches (e.g., interviews, concept-maps, etc.) to measure people's engagement, conceptual understanding, and perceptions about the Earth. McNeal has been awarded over $25 M in external research funding and has published 65 peer reviewed articles in more than 25 different journals.

Lindsay C. Maudlin, PhD, is an Assistant Teaching Professor in the Department of Geological and Atmospheric Sciences at Iowa State University. She teaches several courses related to global climate change and its intersections with society. Additionally, she has researched informal climate change education and used eye tracking to evaluate a decision support system for natural resource management under a changing climate. Lindsay is active in the atmospheric science education research community and is a member of the American Meteorological Society's Committee for the Conference on Education.

Paul R. Armsworth, PhD, is a Professor in the Department of Ecology and Evolutionary Biology and National Institute for Mathematical and Biological Synthesis at the University of Tennessee, Knoxville. An ecological modeler by training, Armsworth's research addresses the integration of ecology and socioeconomic sciences to improve natural resource management and conservation. He has published over 150 peer reviewed papers and book chapters in interdisciplinary outlets as well as disciplinary venues in ecology, evolution, conservation, mathematics, economics and land use policy. In his research, Armsworth collaborates widely with practitioner partners in NGOs, public agencies and for-profit companies.