Research Guided Recommendations for Communicating Uncertainty and Probabilities

Dec 31, 2020

Joseph Ripberger, Andrew Bell, Andrew Fox, Aarika Forney, William Livingston, Cassidy Gaddie, Carol Silva, Hank Jenkins-Smith

NATIONAL INSTITUTE for RISK and RESILIENCE
The UNIVERSITY of OKLAHOMA

Written with support from the NOAA/NWS under award no. NA16OAR4320115.

1. Motivation **2**

2. Methodology **2**

2.1 Study Domain 2

2.2 Identification of Relevant Studies 3

2.2.1 Electronic Search Databases 3

2.2.2 Past Literature Reviews 4

2.2.3 Citation Chains 4

2.3 Extracting Key Information 4

2.4 Evaluating Study Quality 4

3. Results **5**

3.1 Common Topics: Quantity of Evidence 5

3.2 Common Topics: Quality of Evidence 6

3.3 Common Findings 8

3.3.1 Public Understanding and Use of Probability Information in Decision Making 8

Understanding Probability Information 9

Probabilistic vs. Deterministic Information 9

3.3.2 Communicating Probability Information Using Words and Phrases 11

Numeric Translations of Words and Phrases 11

Verbal Directionality and Reference Points 11

Severity and Probability Conflation 12

Choosing Words and Phrases 12

3.3.3 Communicating Probability Information Using Numbers 13

Probabilities as Percentages 13

Probabilities as Frequencies 14

Reference Class Ambiguity 14

Directionality and Reference Points 14

3.3.4 Communicating Probability Information Using Visualizations 15

Visualizations in Non-Weather Domains 15

Visualizations in Weather Domains 16

General Visualization Advice 17

3.3.5 Communicating Probability Information to a Heterogeneous Population 17

4. Recommendations for Forecasters 18

5. Interactive Platform **20**

6. Conclusions and Next Steps **21**

7. References **23**

8. Appendix A: Search Terms **38**

9. Appendix B: Previous Literature Reviews **39**

1. Motivation

Probabilistic forecast information is rapidly proliferating, injecting a new wave of uncertainty into the forecast and warning process. Most scientists agree that this is a positive development, but incorporating probability information into risk communication can be challenging, as probabilities are notoriously difficult to communicate effectively to lay audiences. What does the research literature say about the best way to include probability information in risk communication? What is the evidence base for different communication practices? This project endeavors to address these questions by conducting a "living systematic review" of relevant research from past studies and new studies as they become available.

2. Methodology

A systematic review is a type of literature review that uses a transparent and replicable methodology to identify relevant research from past studies, evaluate results from those studies, and synthesize findings both qualitatively and quantitatively. Historically, systematic reviews have been static; they synthesize the literature at a point in time and become out-of-date almost as soon as they are complete. To prevent this, living systematic reviews are beginning to replace static reviews. Living systematic reviews, like the one we outline in this report, follow the same steps but are updated as new research becomes available. Most systematic reviews, living or static, include some combination of the following steps:

1. Define the study domain
2. Search for and identify relevant studies
3. Extract key topics, questions, methods, and findings from relevant studies
4. Evaluate the quality of relevant studies
5. Analyze and combine the studies to identify common topics, questions, methods, and findings

This review includes two more steps:

6. Assess common findings to develop recommendations to assist forecasters in communicating uncertainty and probabilities
7. Develop a living platform that incorporates new studies and relevant findings into the review as they become available

In the following sections, we use the steps to describe our living review of research literature on the use of probability information in risk communication.

2.1 Study Domain

We focus on research studies that *directly* examine the impact of probability information on protective action decision making, intentions, and behaviors. Most of the studies in the review focus on the "best" or most effective way to communicate probability information. For example, they address questions like: are people more likely to take protective action when probability information is given verbally or numerically? Though extremely important, we do not include studies that *indirectly* examine these relationships by way of implication or suggestion. For example, we do not include studies that explore the relationship between numeracy and risk perceptions, which may have important implications for how people use probability information when making decisions.

2.2 Identification of Relevant Studies

We use a combination of three methodologies to search for and identify relevant studies: (1) electronic search databases; (2) past literature reviews; and (3) citation chains. We use three electronic search databases to search for studies about probabilistic risk/uncertainty communication in the weather and climate domain that use quantitative methodologies. Between all three sources, we were able to identify 327 unique studies that were relevant to the study domain. This list will continue to grow as the review continues.

2.2.1 Electronic Search Databases

In phase one of the search, we use three electronic databases to identify potentially relevant studies that focus on communicating probability information in the weather and climate domains: ProQuest, Web of Science, and EBSCO Academic Search Elite.[1] We restrict the domains at this point in the process to ensure that we are identifying the studies that are most relevant to the audience of this review (see Appendix A for the exact search terms we use). We rely on the next two phases (past reviews and citation chains) to identify potentially relevant studies in adjacent domains (health, insurance, etc.). To begin the review, the first database searches were done on July 29, 2019; repeat searchers were done on September 1, 2020 to update the list of potentially relevant studies. Through these searches, we were able to identify 1,559 potentially relevant studies; 725 of the studies were unique across the three databases.

After identifying potentially relevant studies, two researchers independently screen the title and abstract of each study to ensure that three inclusion criteria are met: (1) the study fits within the study domain (see above); (2) the study reports on new findings from a new research project (e.g., it is not a literature review, essay, or workshop report); and (3) the study uses a reasonably generalizable, transparent, and replicable methodology to conduct the research. In many (but not all) cases, this leads to the exclusion of qualitative studies as most of them do not meet the generalizability criterion. If reviewers do not agree during the screening phase, they review the entire study (text of the article) to see if these criteria are met. Following the first screening phases, the researchers move on to a more in-depth screening phase where they independently review the entire contents of the article to see if it meets the criteria above. If they disagree at this phase, they discuss and come to an agreement. Of the 725 unique articles that we were able to identify at the start of the review, 93 met the inclusion criteria in the first screening and 29 met the criteria following in-depth review of the entire article.

[1] Though it is relatively comprehensive, we do not use Google Scholar because the search results it provides are not replicable.

2.2.2 Past Literature Reviews

In phase two, we use past literature reviews to identify potentially relevant studies. To begin this review, we were able to identify 12 past literature reviews that provide valuable information about potentially relevant studies (see Appendix B for the list of past reviews). Using the two-stage screening methodology we describe above, we came up with 37 new studies that met the inclusion criteria above, bringing the set of relevant studies to 66.

2.2.3 Citation Chains

In phase three, we use citation chains to identify potentially relevant studies. First, we use “backwards” citation chains that include all references IN the relevant studies that we identify in phases one and two. Next, we use “forwards” citation chains that include all references TO the relevant studies that we identify in the first two phases. In all, there were 1,523 unique references IN and 2,279 unique references TO the 66 relevant studies that we were able to identify in the first two phases. Using the two two-stage methodology we describe above, we came up with 255 new studies that met the inclusion criteria above, bringing the set of relevant studies to 327.

Because this is a living systematic review, it is important to reiterate that this is only the beginning of the review. We plan to repeat these phases every few months to make sure that we are including the most up-to-date research.

2.3 Extracting Key Information

Following identification, we carefully scan all relevant studies, extract key information, and store it in an electronic spreadsheet. In addition to basic bibliographic information, we note relevant research questions, independent (treatment) and outcome variables, research methodologies (i.e., survey experiment), information about research subjects, domains of study (i.e., weather, health, etc.), and primary findings.

2.4 Evaluating Study Quality

While extracting key information about each study, we also assess the quality of the studies. There are many ways to assess quality; we use three indicators of validity that are important and commonly known in the social and behavioral sciences: (1) external validity (are the results generalizable to the population of interest?); (2) internal validity (are we sure that variation in x causes variation in y?); and (3) domain validity (how relevant is the study domain to weather hazards and forecasting?). Each dimension is independently given a score by two researchers on a three point scale (1 = low; 2 = medium; 3 = high). We use the mean value of these scores to measure validity along each dimension and the overall validity of each study.

3. Results

3.1 Common Topics: Quantity of Evidence

We were able to identify 5 primary research topics and 14 secondary topics in the set of 327 studies that we are using to begin this review (see below for a description of the topics). While this list of topics is likely to grow as the review continues, an examination of the relative frequency of each topic provides a snapshot of the literature at this point in time. It provides valuable information about the *quantity* of evidence we were able to identify in each topic area.

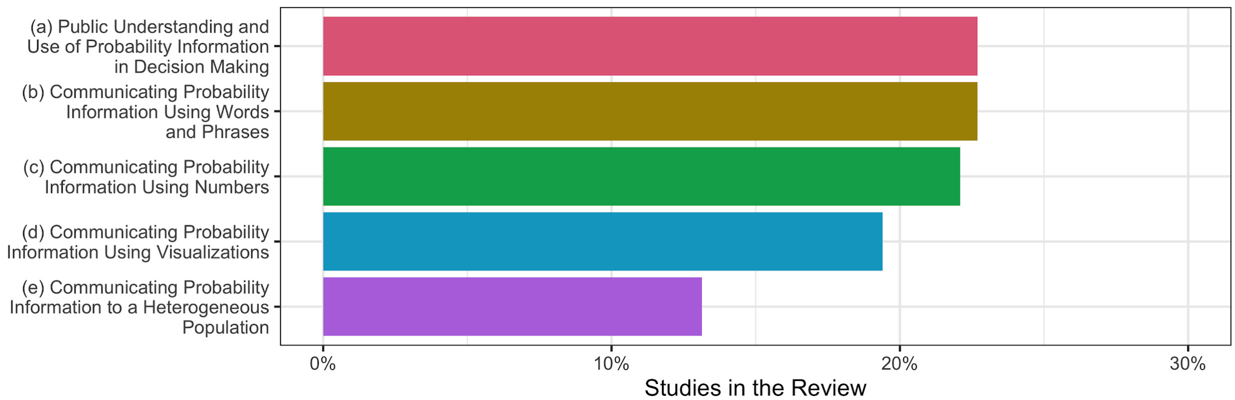

Fig. 1: Frequency of Relevant Studies by Primary Topic

Fig. 1 provides this information by showing the percentage of studies that address the primary topics. Note that the topics are not mutually exclusive; many studies address more than one topic. As Fig. 1 indicates, the most common topics in the current set of relevant studies are *(a) public understanding and use of probability information in decision making* and *(b) communicating probability information using words and phrases*. The least common topic in the current set of relevant studies is *(e) communicating probability information to a heterogeneous population*. The discrepancy between the two topics indicates that we know a lot about how people use probability information to make decisions and the extent to which words and phrases facilitate this process, whereas we know relatively little about how different types of people use and interpret different types of probability information.

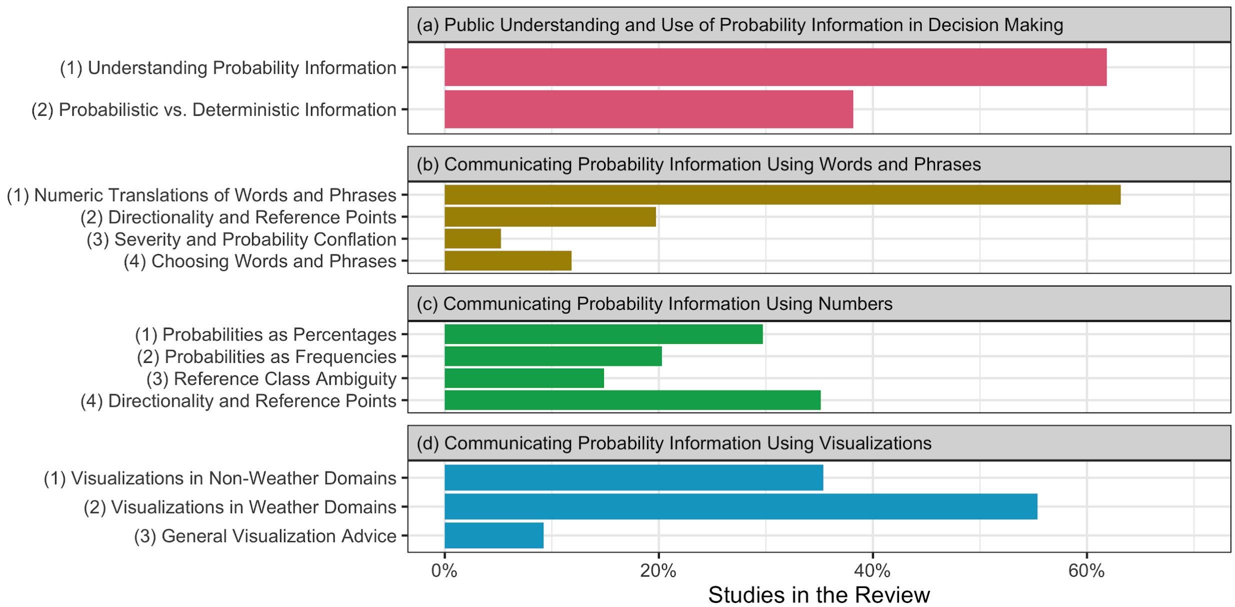

Fig. 2: Frequency of Relevant Studies by Secondary Topic

Fig. 2 provides more information by showing the percentage of studies within each primary topic that address secondary topics of relevance. For example, the figure shows that studies on the *(b) communication of probability information using words and phrases* frequently address *(1) numeric translations of words and phrases*, but rarely address *(3) severity and probability conflation*. Likewise, studies on the *(c) communication of probability information using numbers* are more likely to focus on *(1) probabilities as percentages* than *(2) probabilities as frequencies*. Again, the discrepancies we observe within each topic area provide valuable data on the amount of evidence we have on each of the secondary topics.

3.2 Common Topics: Quality of Evidence

Information about *quantity* provides one metric for assessing the strength of evidence in each of the topic areas. Information about *quality* of evidence provides a second metric. As we note above, we assess the quality of evidence from each study in the review by evaluating the external, internal, and domain validity of the study. When we average these across topics, we can discriminate between topics with relatively high quality evidence and topics with low quality evidence. As with quantity of evidence, these averages will change as the review continues, but they provide valuable information about the state of the research literature at this point in time.

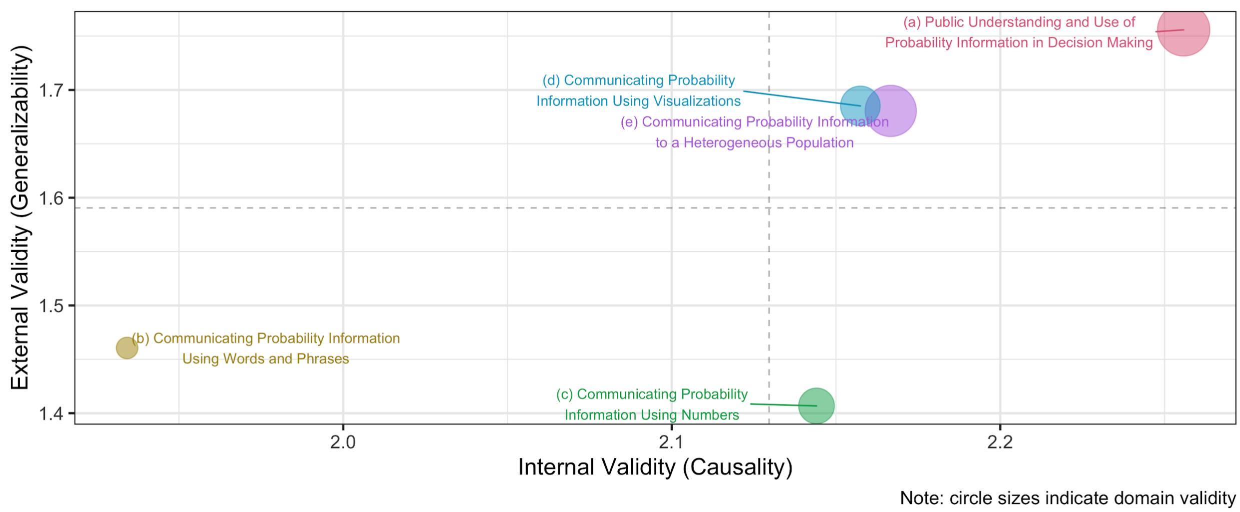

Fig. 3: Validity of Relevant Studies by Primary Topic

Fig. 3 provides this information by showing the mean validity of studies that address the primary topics. Note that the sizes of the points in the plot reflect mean domain validity; relatively large points indicate high domain validity, small points indicate low domain validity. The lines with dashes indicate overall mean validity scores; on average, studies on topics to the *right* of the line have more internal validity than studies to the *left* of the line and studies *above* the line have more external validity than studies on topics *below* the line. Given this orientation, the relatively large point in the top-right quadrant indicates that, on average, studies on *(a) public understanding and use of probability information in decision making* have extremely high validity. By comparison, the relatively small point in the bottom-left quadrant indicates that studies on *(b) communicating probability information using words and phrases* generally have low validity. Interestingly, studies on *(c) communicating probability information using numbers* typically have more internal but less external validity than studies that focus on the use of words in phrases -- this is because these studies often use experiments to identify causality, but the subjects of the experiments are rarely representative of the US population.

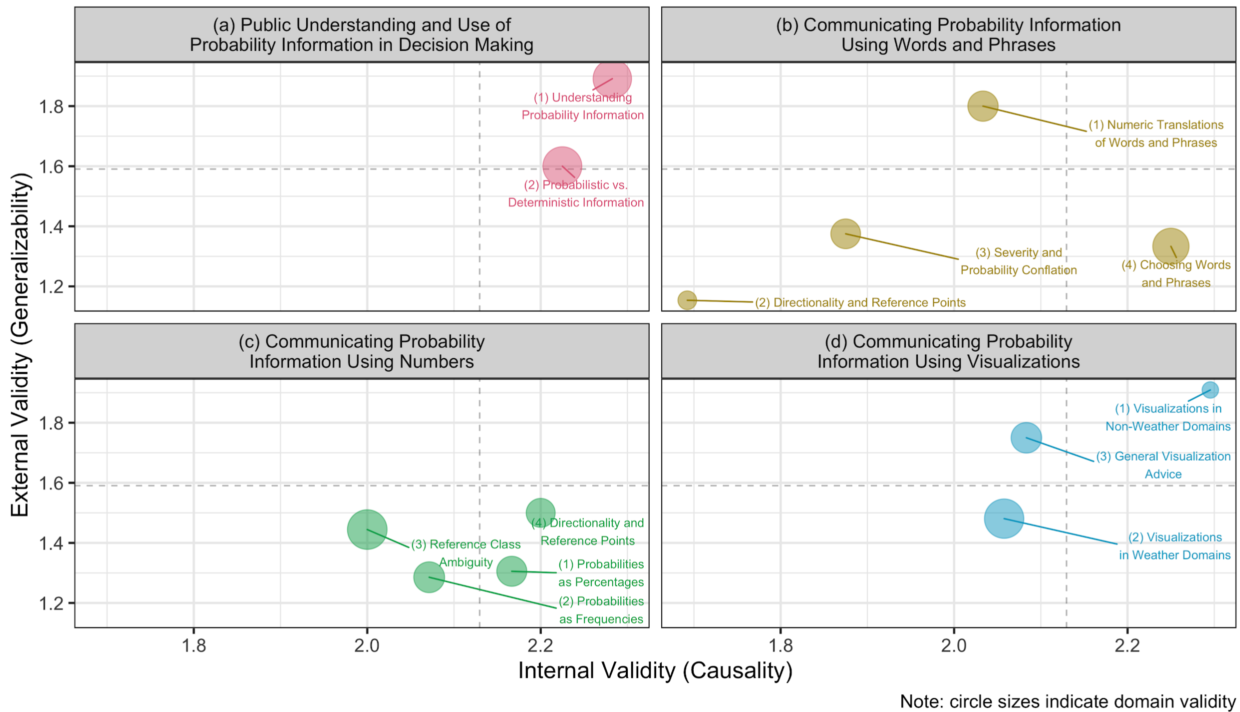

Fig. 4: Validity of Relevant Studies by Secondary Topic

Fig. 4 provides more information by plotting the average validity of studies by secondary topic within each of the primary topics. The plots illustrate, for example, that studies on *(1) understanding probability information* and *(2) probabilistic vs. deterministic information* have especially high levels of validity, whereas studies on *(2) directionality and reference points* and *(3) severity and probability conflation* (when using words and phrases) typically have low levels of validity. If quality is the metric, these findings suggest that we know a lot about the first two areas of study, whereas more research is necessary on the latter two topics of inquiry.

3.3 Common Findings

We were able to identify more than 100 unique findings in the set of 327 studies that we use to begin this review. It is not possible to independently describe all of these findings within the confines of a single report. We therefore attempt to synthesize and summarize as many findings as possible in the sections below.

3.3.1 Public Understanding and Use of Probability Information in Decision Making

Before we can think about the best way to communicate probability information, we must first consider whether probability information is even useful to members of the public. The first topic in the review deals primarily with this question, and encompasses all studies examining how people understand and think about probabilities, or comparing the usefulness of probabilistic forecasts to deterministic forecasts.

Understanding Probability Information

Does the public actually understand, use, and benefit from probability information, or are they better off with deterministic statements? Some forecasters express a desire to "boil down" complex probability information to a deterministic point forecast for fear of confusing members of the public (Pappenberger et al. 2013). However, the evidence in this review suggests quite strongly that those fears are unfounded. Nearly all of the studies we review indicate that people make better decisions, have higher trust in information, and/or display a greater understanding of forecast information when shown a probabilistic forecast instead of a deterministic one (Ash et al. 2014; Bolton and Katok 2018; Grounds and Joslyn 2018; Grounds et al. 2017; Joslyn and LeClerc 2012; Joslyn and LeClerc 2016; Joslyn and Demnitz 2019; Joslyn et al. 2007; LeClerc and Joslyn 2012; Marimo et al. 2015; Miran et al. 2018; Nadav-Greenberg and Joslyn 2009; Roulston and Kaplan 2009; Roulston et al. 2006; Joslyn and Grounds 2015). The majority of these studies are firmly situated in the context of weather; however, it is important to note that both experts and the public sometimes have difficulty interpreting probability information, and different ways of communicating risk can lead to variation in understanding (Bramwell et al. 2006). Taken together, the studies in this review indicate a very strong and consistent pattern of results throughout the literature -- simple probabilistic forecasts outperform deterministic ones in almost all situations and by almost all metrics. This does not mean that the average forecast end-user is a competent statistician; in fact, many of these studies emphasize the need to make probabilistic forecasts as straightforward and easy to understand as possible in order to avoid "information overload" (Durbach and Stewart 2011). It is also important to note that these findings refer to probability information beyond a general acknowledgement of epistemological uncertainty (e.g. "this forecast is based on estimates; it is impossible to ever know for sure what will happen"), as these types of overly broad statements can undermine trust in forecasts (Howe et al. 2019). This research should, however, reassure those who worry that the public will be confused by anything other than a deterministic forecast.

Probabilistic vs. Deterministic Information

Many studies also examine *how* people understand and interpret probability information. Often, these studies simply give subjects a Probability of Precipitation Forecast (PoP) and ask them to interpret it in order to assess their level of understanding. Gigerenzer et al. (2005), Morss et al. (2008), Sink (1995), Zabini (2015), and Abraham et al. (2015) all show, to various degrees, that a majority or a substantial proportion of the public is unable to give the correct interpretation of a PoP forecast, generally considered to be something like "it will rain somewhere in the forecast area on X% of days like today" (Gigerenzer et al. 2005; Morss et al. 2008; Sink 1995; Zabini 2015; Abraham et al. 2015). However, Juanchich and Sirota (2018) argue that previous studies use a cumbersome "correct" answer, and that "X% of simulations predict rain in the forecast area" is a more "fluent" and more easily understood response category; when using this "correct" answer, the vast majority are able to give the correct PoP interpretation (Juanchich and Sirota 2018). Another prior study, Murphy et al. (1980), argues something similar -- that "people do not have trouble understanding what "30% chance" means, but… they do have trouble understanding exactly what the probability refers to in this kind of forecast" (Murphy et al. 1980). These findings should reassure forecasters that the public can correctly understand probabilities, but should underscore the need to explain the *events* the forecast refers to in an intuitive and clear way.

Multiple studies also offer findings about the process by which members of the public think through probabilistic risk information. For instance, a group of studies indicate that overwhelming majorities of the public intuitively infer uncertainty even when given a deterministic forecast (Savelli and Joslyn 2012; Joslyn and Savelli 2010; Morss et al. 2008; 2010). Moreover, inclusion of uncertainty can lead to increased worry but can also be mitigated through the use of textual and visual formats (Han et al. 2011). These findings suggest that people think about forecasted events in probabilistic terms even when not explicitly told to do so. Several more studies find that higher probabilities (regardless of context or direction) may lead people to view a forecast as more accurate; for instance, the same forecast would likely be taken to be more accurate if it reported a 70% chance of sun rather than a 30% chance of rain (Bagchi and Ince 2016; Lohre et al. 2019; Juanchich and Sirota 2017). In a similar vein, the research in this review suggests that people consistently misinterpret confidence intervals and forecast periods. For instance, two studies find that both experts and non-experts implicitly interpret forecasted events as being more likely towards the end of the forecast period (e.g. if there were an X% chance that a given event would occur sometime in a given week, people will, on average, perceive that the event is more likely to happen on Friday than on Monday) (Doyle et al. 2014; McClure et al. 2015), and one study finds that a significant proportion of the public is unsure of how to understand the distribution of possible outcomes denoted by confidence intervals (Dieckmannn et al. 2015). These findings underscore the need to clarify what forecast periods and confidence intervals mean in the context of a given forecast rather than assuming that will be clear to the audience.

Finally, a few studies also identify some common ways in which audiences can be led to misinterpret or dismiss probability information in risk messages. Perhaps the most significant of these challenges is that of motivated reasoning, which is the tendency to interpret new information in a way that supports preexisting beliefs. Four studies in the review explicitly focus on the effect of motivated reasoning on how people interpret probability information; all indicate that probability risk information, especially information about politically sensitive topics like gun control or climate change, is susceptible to misinterpretation when it contradicts preexisting beliefs (Budescu et al. 2012; Dieckmann et al. 2017; Piercey 2009; Nurse and Grant 2020). Interestingly, more numerate people seem to be more susceptible to these effects (Nurse and Grant 2020), and using verbal probability expressions seems to encourage more motivated reasoning than using visual or numeric expressions (Budescu et al. 2012; Dieckmann et al. 2017; Piercey 2009). Another common way in which probability information can become "distorted" is what Hohle, and Teigen (2015) call the "trend effect" -- in short, they and others in the review show that non-experts interpret recent forecasts in light of what previous forecasts have said. A "moderate" risk, for instance, will cause more worry if it has been upgraded from a "low" risk than if it has been downgraded from a "high" risk (Hohle, and Teigen 2015; Løhre 2018).

3.3.2 Communicating Probability Information Using Words and Phrases

Another significant portion of the research in this review deals with the use of words and phrases (in place of or in addition to numbers) to communicate probability information. Experts and non-experts both routinely use verbal expressions such as “unlikely” or “a good chance” to describe uncertainty in all domains, but the practice is particularly common in the realm of weather and climate -- for instance, the Intergovernmental Panel on Climate Change (IPCC)’s annual reports use primarily verbal risk expressions in an attempt to better communicate forecasts to non-experts. However, despite their popularity, it is often unclear how best to use them and how these types of communications differ from numeric or visual depictions of risk; the research discussed in this section helps to answer those questions.

Numeric Translations of Words and Phrases

The first core finding in this area of the review is very simple: there is strong evidence that risk communicators should *always* include a numeric “translation” for any verbal probability expressions used, and that translation should appear directly in or next to the verbal expression itself (Carey et al. 2018; Connelly and Knurth 1998; Dorval et al. 2013; Fortin et al. 2001; Hill et al. 2010; Zabini et al. 2015; Wintle et al. 2019). For example, a verbal expression like “severe thunderstorms are possible this evening” would be more effectively rephrased as “severe thunderstorms are possible (20% to 30% chance) this evening” (Budescu et al. 2014). Explicit statements of the upper and lower bound (e.g. 0-33%) implied by an expression (e.g. “likely” or “unlikely”) improve accuracy of interpretation vs. a statement alone (Harris et al. 2017). This is important not only because it helps people to correctly interpret the meaning of a forecast, but also because people generally prefer mixed formats (e.g. a numeric probability and a verbal probability expression together, or a number and a visualization) to singular ones (Carey et al. 2018; Connelly and Knurth 1998; Dorval et al. 2013; Fortin et al. 2001; Hill et al. 2010; Sink 1995; Zabini et al. 2015). Members of the public are able to demonstrate basic understanding of probabilistic forecasts; however, uncertainty is best communicated with a range of estimates rather than single points and through combined use of numeric and verbal expressions to meet the needs of heterogeneous audiences (Kox et al. 2015). Translations and numerical probabilities can sometimes lead to misinterpretation due to a number of factors related to both forecasters and audiences, such as use of correct terms, preferences, and numeracy (Sink 1995). Furthermore, translations are important because less numerate people tend to focus on narrative evidence when evaluating risk communications (the context, their perceptions regarding the likelihood of comparable events, etc.), while more numerate people tend to focus on the numeric probability of the risk (Budescu et al. 2009; Dieckmann et al. 2009; Budescu et al. 2012; Juanchich et al 2013; Mandel 2015).

Verbal Directionality and Reference Points

The next core finding in this research addresses the importance of “directionality” and “reference points” in verbal probability statements. “Directionality” refers to whether a statement emphasizes the chance that something *will not* happen (e.g. “it is likely that the hurricane will miss our town”) or the chance that something *will* happen (e.g. “it is unlikely that the hurricane will hit our town”). Findings from these studies suggest that forecasters should be mindful of these concepts when crafting verbal probability statements, as they can have unintended consequences on interpretation. It should be noted that there is no universal "best" choice to use when it comes to directionality and contextual factors can affect understanding, highlighting the ability for forecasters to accidentally imply that their forecasts are higher or lower than intended, thereby indirectly influencing how people interpret the communication. Generally, positive statements seem to draw attention to the possibility of an event happening and cause people to interpret such statements as more likely, while negative statements tend to have the opposite effect, promoting risk aversion by encouraging people to focus on the possibility of negative events, even if they are unlikely (Honda and Yamagishi 2006; 2009; 2017; McKenzie and Nelson 2003; Teigen and Brun 1995; 1999; 2000; 2003, Wallsten et al. 1986; Budescu et al. 2003). Negative modifiers (i.e. "not likely", "improbable", etc.) tend to be interpreted as more qualifying and less definitive than positive verbal modifiers (i.e. "very possible", "quite likely", etc.) (Reyna 1981). Finally, research indicates that the construction of a probability statement affects how people interpret it as well; independent of the directionality of the phrase, people generally interpret verbal probability statements as more likely when they describe the chance of failure (or an undesirable outcome) as opposed to the chance of success (Juanchich et al. 2013; Mandel 2015).

Severity and Probability Conflation

Another core finding of the literature regarding verbal probability expressions pertains to the "severity effect", which is the tendency for people to implicitly interpret verbal probability expressions as more likely when they describe more severe or undesirable outcomes (Bonnefon and Villejoubert 2006; Fischer and Jungermann 1996; Harris and Corner 2011; Weber and Hilton 1990). For example, someone who interprets a "slight chance" of rain showers to mean a 1-5% chance will likely interpret a "slight chance" of a hurricane to mean something closer to a 10-15% chance. This is important for forecasters to consider when using verbal probability statements, as it may suggest different interpretations of the same words and phrases, depending on the situation.

Choosing Words and Phrases

Finally, the studies in the review provide a few core findings regarding word choice. For instance, when deciding whether to use a word like "can" or "will", be aware of the "extremity effect": when shown a probability and asked what "can" happen, people tend to focus on the most extreme possible values, and when asked what "will" happen, they tend to focus on the more likely scenarios (Teigen and Filkuková 2013; Teigen et al. 2018; Teigen et al. 2014). If focusing on more remote possibilities would be counterproductive to the goal of communications, forecasters should be wary of what word choice implies about certainties and risks. In a similar vein, Teigen, Juanchich, and Riege (2013) find that people often use and interpret words like "improbable" to refer to events that aren't just unlikely (something like a 10-20% chance, for instance), but nearly impossible (closer to a ~1% chance, for instance). Often, "improbable" is implicitly understood to refer to events that have not happened yet but have a small chance of happening in the future, even when experts have another definition in mind. Clarifying such terms and providing explicit numeric "translations" helps to reduce these misunderstandings (Teigen, Juanchich, and Riege 2013). Lastly, research strongly indicates that forecesters should vague verbal probability terms (such as “it is possible” or “there is a chance”), as they can be particularly problematic in communication due to variable interpretation (Fillenbaum et al. 1991; Reyna 1981; Lenhardt et al. 2020). In sum, words and phrases play an important role in the communication of probability information. As a result, forecasters and audiences would benefit from careful consideration of translations, directionality, severity, and word choice to ensure clear communication.

3.3.3 Communicating Probability Information Using Numbers

Both experts and non-experts frequently present probabilistic risk information numerically using a variety of formats, such as percentages and frequencies. Format is also directly related to how information is framed. Framing enables individuals and society to make sense of the world around them by providing a lens through which to interpret events, ideas, and risks (Goffman 1974; Kühberger 1998; Miller 2000). As a result, framing can have a direct influence on the way members of the public understand risk and interpret probability information (Goffman 1974; Tversky and Kahneman 1981; Benford and Snow 2000; Kitzinger 2007).

Probabilities as Percentages

Numeric probabilities are most commonly expressed as a percentage (e.g. “a 30% chance of rain”) or as a frequency (e.g. “a 3 in 10 chance of rain”). In general, the use of probability (percentage) formats has been observed to improve the accuracy of risk comprehension and encourage protective action over the use of verbal probability statements alone (Fuller et al. 2001; Fuller et al. 2002; Koehler 2001). Some evidence suggests, however, that when comparing between probability and frequency formats, there are no consistent and significant benefits to one over the other (Hendrickx et al. 1989; Joslyn et al. 2009i; Neace et al. 2008; Knapp et al. 2016; Evans et al. 2000; Ruiz et al. 2013; Strathie et al. 2017). Instead, the research in this review suggests that different circumstances call for different formats: advantageous use of frequency and percentage formats depends on the given task at hand and the context of the risk (Cuite et al. 2008; Knapp et al. 2009 ; Sinayev et al. 2015; Wallsten et al. 1986).

Cuite et al. (2008) investigate the effects of format on interpretation of probabilistic risk information and show that percentage formats have high overall accuracy rates; however, they also noted that communicating risk information of this nature also requires awareness that not all formats perform well in every situation or type of hazard. Percentage probabilities describing rare, low-probability events are particularly prone to misinterpretation, for example (Cuite et al., 2008). Absolute risk formats can communicate higher levels of urgency and risk perceptions (Galesic et al. 2009; Zikmund-Fisher et al. 2008; Ilic et al. 2012); however, while absolute risk formats can produce greater accuracy, users among the general public might find such presentation of information overwhelming or even confusing (Knapp et al. 2010).

Probabilities as Frequencies

Frequency formats, such as “1 in X” and “X in NX” formats, have also been found to influence decision making and risk comprehension (Grimes and Snively 1999; Pighin et al. 2015; Oudhoff and Timmermans 2015; Denes-Raj et al. 1995; Bell and Tobin 2007; Carey et al. 2018; Pighin et al. 2011). For example, Pighin et al. (2015) assess the effects of format ("1 in X" or "N in NX") on pregnant women’s interpretations of risk information. They find that women had elevated perceptions of risk when shown the 1 in X format as compared to the N in NX format. Interestingly, this was still true despite no observed effect on comprehension (Pighin et al. 2015). Similarly, Oudhoff and Timmermans (2015) examine format and graphic effects on decisions to engage in lotteries. They note that the 1-in-X format yielded higher perceived likelihoods and appeared to be the easiest format to interpret; however, graphs primarily affect perceptions of likelihood among people with lower numeracy (Oudhoff and Timmermans 2015). Finally, Denes-Raj et al. (1995) investigate the effects of using a “1 in X” format compared to “N in NX” format in order to better understand how ratio bias plays a role in decision making. Their research suggests that although ratio bias is common, it is not ubiquitous across different situations (Denes-Raj et al. 1995).

Reference Class Ambiguity

Under some circumstances, the benefits of using percentage probability formats in place of frequency formats are less evident (Hendrickx et al. 1989; Joslyn et al. 2009; Neace et al. 2008 ; Knapp et al. 2016; Evans et al. 2000 ; Ruiz et al. 2013; Strathie et al. 2017). For instance, Evans et al. (2000) observe that formats, whether frequencies or percentages, have little effect on understanding, unless they are also presented in a way that “facilitated construction of a set inclusion mental model” (p. 197). Gigerenzer et al. (2005) argue that one way to provide such a definition is through the inclusion of reference class, or the class of events to which a probability refers. Rather than simply stating, “There is a 30% chance of rain tomorrow,” a forecast that includes reference class would be written, “There is a 30% chance of rain on days like tomorrow. This means that on 3 out of 10 days like tomorrow, it will rain” (Gigerenzer et al. 2005). Juanchich and Sirota (2018) find that specifying the reference class can increase correct interpretation by reducing misunderstanding and ambiguity in the way forecasts are presented. Neace et al. (2008) present evidence that supports using frequency formats to reduce errors, they also demonstrated that when controlling for reference class, this effect largely disappears, leaving little difference between frequency-based and probability-based formats.

Directionality and Reference Points

Directionality of a statement can also affect the use of probability information (Honda and Yamagishi 2017). Positive probability statements are often more effective at spurring people to take precautionary actions (Teigen and Brun 1999), increase perceived benefits (Flugstad and Windschitl 2003; Kupor and Laurin 2020), and produce higher risk perceptions, especially when expressed in frequency formats (Wu et al. 2018). Similarly, Garcia-Retamero and Cokely (2011) find that gain-based framing tends to increase preventive behaviors.

Wernsted et al. (2019) find similar effects in severe weather scenarios, observing that gain-based frames resulted in more risk-averse behavior than loss-based framing. Gain-based framing led nearly two-thirds of emergency managers to make decisions that avoided risks (Wernsted et al. 2019).

When considering the effects of positive and negative framing on willingness to take risks, the empirical evidence suggests that messages with negative frames can discourage individuals from taking actions they perceive as risky (Morton et al. 2011; Levin et al. 1998; Keren and Gerritsen 1999). Negative frames in general, and loss-based frames in particular, have also been found to increase risk perception and encourage detection-related behavior (Grimm et al. n.d.; Banks et al. 1995; Chua et al. 2006; Garcia-Retamero and Cokely 2011; Teigen and Brun 2003). When presented as more severe, Patt and Schrag (2003) observe that forecasts increase the perceived likelihood of the risk. Similarly, the use of time uncertain formats for presenting future event probabilities increases risk perceptions and endorsement of protective decisions (Ballard and Lewandowsky 2015).

However, more recent research suggests that loss-based framing can often interact with higher levels of uncertainty to decrease individual intent to take action in longer time frame events, such as climate change (Morton et al. 2011) and possibly lead recipients to misinterpret forecasts (Gibson et al. 2013). Similarly, Kuhn (1997) finds that negative framing often influences individuals to prefer ambiguous probabilities, using terms like "about a 20% chance," which can reduce the accuracy of comprehension. Moreover, Morton et al. (2011) shows that negative framing, when presented with higher levels of uncertainty, tends to decrease intent to engage in protective behaviors.

As a result, negative phrasing can sometimes have a deleterious effect on lay judgment and interpretation of probability expressions (Smithson et al. 2012) as well as a detrimental effect on decision making when aimed at less numerate audiences and in situations involving specific risks with high uncertainty (Kuhn 1997 ; Morton et al. 2011 ; Peters and Levin 2008; Armstrong et al. 2002; Peters et al. 2006). These effects can be mitigated through the use of experience-based formats to reduce the tendency for individuals to overrate rare risks and rely on emotions for decision-making (Tyszka and Sawicki 2011).

3.3.4 Communicating Probability Information Using Visualizations

Visualizations often provide an effective way to communicate probability information to groups who may have difficulty with probabilities expressed as numbers or words, even if for no other reason than because people want them and, subjectively, find them useful (Johnson and Slovic 1995; 1998). Use of visuals has also been shown to be more effective in communicating probability information with younger audiences (Ulph et al. 2009), and less numerate audiences (Okan et al. 2015), making them a crucial tool for risk communicators in any domain.

Visualizations in Non-Weather Domains

Certain types of probability visualizations are common in the weather domain, but many are not; icon arrays, pictographs, risk ladders, and survival curves, for example, are common in the fields of medicine and epidemiology, but are rarely seen in weather forecasting. Icon array visualizations can be used to increase risk comprehension and understanding of probability information. Many studies show that icon arrays can increase understanding and risk avoidance actions (Ancker et al. 2011; Tubau et al. 2019; Galesic et al. 2009; Garcia-Retamero et al. 2010; Garcia-Retamero and Galesic 2009; Keller and Siegrist 2009; Schirillo and Stone 2005; Garcia-Retamero and Cokely 2014; Garcia-Retamero and Galesic 2010; Stone et al. 1997; Taylor et al. 2018; Witteman et al. 2014; Zikmund-Fisher et al. 2014; Zikmund-Fisher et al. 2011). Icon arrays can also help to communicate risk information in a way that shows exact percentages while simultaneously conveying "gist" impressions (Hess et al. 2011). Research indicates that pictographs can be helpful in understanding probability information as well. Pictographs can be helpful when aiming to lower risk perceptions (Dowen et al. 2017; Hawley et al. 2008; Keller and Siegrist 2009; Leonhardt and Keller 2018). Pictographs help to better convey individual units (people) within the at-risk population. Risk ladders and other visualizations can help to improve people's comprehension of information about very small or long-term risks but should be used with caution because individuals with low numeracy may have difficulty understanding ladders without comparative information (Keller et al. 2009). Survival curves are common in medical and epidemiological settings; however, research indicates that audiences are often unfamiliar with these representations, leading to misinterpretation of probability information without specific instructions about how to interpret them (Armstrong et al. 2001; Mazur and Hickam 1990).

Visualizations in Weather Domains

Pictographs and icon arrays, however, are relatively rare in the domain of weather due to the type of risks involved; instead, weather-related probability visualizations typically involve displaying risk on a map. In general, it should be noted that people understand basic probability information about forecasts when presented with a map (Wu et al. 2014). Gerst et al. (2020) find that using simplified maps (e.g., legends, contour lines, white space) can improve overall understanding of probability information. Maps communicate risk to audiences because they communicate geographic proximity to threats, thereby encouraging calibration of risk perceptions (Roth 2009). Color and size-based visualizations allow for quicker but less thought-out decisions, while texture and icon-based depictions require more time and deliberation (Cheong et al. 2019; Seipel and Lim 2017; Leitner and Buttenfield 2000; Miran et al. 2017; Tak and Toet 2014.; Miran et al. 2016; Retchless and Brewer 2016; Sherman-Morris et al. 2015). In addition to colors, track lines have been found to increase levels of concern among members of the public with little effect on the interpretation of probability information (Meyer et al. 2013; Newman and Scholl 2012; Padilla et al. 2017; Ruginski et al. 2016; Van Pelt et al. 2015; Wu et al 2014).

In addition to maps, other types of visualizations have been used to improve understanding of probability information in weather forecasts, including ensemble plots (Toet et al. 2019) and interactive formats (Hogarth and Soyer 2015; Natter and Berry 2005). For example, the visualization of a confidence interval can influence how people interpret a forecast. People understand smaller confidence intervals to mean that a forecast is more technologically advanced and more certain (Teigen et al. 2018). Additionally, predictive interval graphics can encourage protective action (S. Joslyn, Nemec, and Savelli 2013). Use of visualizations, however, can often depend on the purpose of communication. In some instances, the use of deterministic visualizations can increase protective action, while the use of probabilistic visualizations can better communicate changes in risk (Ash et al. 2014; Baker 1995; Marimo et al. 2015).

General Visualization Advice

There is no "best visualization format" for conveying risk, so it is crucial that forecasters think carefully about the situation at hand and the audience they intend to reach (Barnes et al. 2016; Bisantz et al. 2005; Etnel et al. 2020; Garcia-Retamero and Dhami 2013; Kreye et al. 2012; Lorenz et al. 2015; Sanyal et al. 2009). Experts should consider what their audience prefers, what research suggests works best for the situation, and the magnitude/danger of the risks involved. While there is no "best" format, some of the studies in this review suggest that there are some visualizations/practices that should generally be avoided. For instance, multiple studies indicate that communicators should not use bar charts to communicate probability information, as they can increase misunderstanding (Correll and Gleicher 2014 ; Dieckman et al. 2015; Schapira et al. 2006; Newman and Scholl 2012). Similarly, the use of survival curves should be avoided with audiences unfamiliar with these representations as they tend to lead to misinterpreting probability information (Armstrong et al. 2001; Mazur and Hickam 1990).

With visualizations, there are some issues with presentation that influence how the audience may interpret information, regardless of the visual format used. To combat some of these issues experts should use foreground displays to increase risk perceptions but with the knowledge that effects are contingent on depicted probability sizes, labels, and risk reduction levels associated with protective action (Okan et al. 2020; Okan et al. 2018; Stone et al. 2017). More importantly, studies indicate that communicators ought to include descriptive labels with graphical displays to improve risk understanding among people with low numeracy, who sometimes struggle to follow complex visualizations (Okan et al. 2015).

3.3.5 Communicating Probability Information to a Heterogeneous Population

Lastly, many of the studies in the review include findings that, while important, are overly specific or difficult to generalize and use for forecasters. The general consensus of these findings, though, is clear: efforts to communicate probability information must also consider the heterogeneous nature of audiences and the different contexts and biases they entail. For instance, when it comes to both members of the public and experts, individuals sometimes have difficulty accurately interpreting forecasts that include probability information due to a number of different factors, such as context (Kim et al. 2014; Klockow-McClain et al. 2020; Morss et al. 2010; Hohle, and Teigen 2015; Løhre 2018; Windschitl and Weber 1999), motivated reasoning (Dieckmann et al. 2017), and numeracy (Kong et al. 1986; Bramwell et al. 2006; Harris et al. 2013; Rinne and Mazzocco 2013; Juanchich and Sirota 2016). Forecasters play a critical role in setting the context for communicating probability information through their use of language (Connelly and Knurth 1998; Franic and Pathak 2000; Kox et al. 2015; Pappenberger et al. 2013). The way in which forecasters frame messages can influence how audiences interpret risks (Harris et al. 2009; Keller et al. 2006; Wilson et al. 2019). There is a real risk of inducing audiences to overestimate risk when describing small or long-term risk using a "once in X years" expressions (Grounds et al. 2018; Bell and Tobin 2007), time uncertain (e.g. "X will happen within the next 10-30 years") (Ballard and Lewandowsky 2015), and longer time frames (Keller et al. 2006; Doyle et al. 2014; McClure et al. 2015; Morss et al. 2016). Audiences are also prone to motivated reasoning when presented with probability information and uncertainty (Savelli and Joslyn 2012; Joslyn and Savelli 2010; Morss et al. 2008; 2010; Dieckmann et al. 2009; Rabinovich and Morton 2012). Although uncertainty often elicits more risk-averse behavior (Ramos et al. 2013), individuals have a tendency to engage in additivity neglect (e.g. responses totaling more than 100%) (Riege and Teigen 2013), anchoring (Losee et al. 2017), and favoring higher absolute numbers (Denes-Raj and Epstein 1994; Maglio and Polman 2016). Lastly, less numerate people tend to focus on narrative evidence when evaluating risk communications (the context, their perceptions regarding the likelihood of comparable events, etc.), while more numerate people tend to focus on the stated probability of the risk with greater accuracy (Dieckmann et al. 2009; Gardner et al. 2011).

4. Recommendations for Forecasters

This is the most difficult and subjective step in the review. In this step, we work as a team to translate core findings into actionable recommendations. This is challenging because we have to decide which of the core findings are certain and relevant enough to support a recommendation. Relevancy is the primary challenge. Many studies in this review advance basic science but do not relate the findings to specific communication challenges. While valuable, it can be difficult to translate these findings into actionable advice for risk communicators. In addition, many studies in the review provide actionable advice on how to communicate probability in domains like healthcare. Again, the insight from these studies can be extremely valuable but difficult to translate into actionable advice for risk communicators who work in extreme weather/climate domains. This is where the subjectivity enters into the process. We pick the core findings that we believe are most relevant to risk communicators in extreme weather/climate domains below.

In Table 1 we list the core findings that we believe are certain and relevant enough to support tangible recommendations. We also include a communication example and example study for each recommendation. While we have extensive experience working with risk communicators in this domain, we cannot know or define everything that might be relevant. We therefore urge risk communicators to carefully examine core findings on their own to see if there might be something of relevance that we are missing in the list of recommendations.

Table 1: Recommendations for Forecasters

Recommendation	Example	Exemplary Studies
Use probability information in place of deterministic statements in forecasts.	Replace: These storms will cause heavy downpours and flooding. With: There is an extremely high (90%) chance that these storms will cause heavy downpours and flooding.	Joslyn & LeClerc (2012); Grounds & Joslyn (2018)

Use probability ranges (predictive intervals) to emphasize uncertainty when point estimates are not available or appropriate; wide ranges indicate more uncertainty.	Replace: The forecast is rapidly evolving, but there is a slight (15%) chance that we will see more than 10 inches of snow in the metro area tomorrow morning. With: The forecast is rapidly evolving, but there is a slight (10% to 20%) chance that we will see more than 10 inches of snow in the metro area tomorrow morning.	Grounds et al (2017); Løhre et al. (2019)
Include numeric translations next to words/phrases that indicate probability information.	Replace: Severe thunderstorms are possible this evening. With: Severe thunderstorms are possible (20% to 30% chance) this evening.	Budescu et al. (2014)
If comprehension of probability information is especially important, use numeric probabilities alone or first (before words/phrases).	Replace: Severe thunderstorms are possible (20% to 30% chance) this evening. With: There is a 20% to 30% chance of thunderstorms this evening.	Jenkins et al. (2019)
When using words and phrases to communicate probability information, include rank adjectives (like low, medium, and high) to indicate the magnitude of the probability; this is especially important if numeric translations are not available.	Replace: There is a chance of snow and ice this morning along I-75. With: There is a low/medium/high chance of snow and ice this morning along I-75.	Lenhardt et al. (2020)
Use probability (percentage) formats when possible; frequency (fraction) formats can be effective, but they can also generate confusion.	Replace: There is a 1 in 10 chance that this storm will produce a tornado. With: There is a 10% chance that this storm will produce a tornado.	Cuite et al. (2008); Joslyn & Nichols (2009)
When using frequency (fraction) formats, use 1 in X formats in place of X and NX formats.	Replace: Recent models indicate that there is a 20 in 100 chance that this storm will cause significant storm surge in the New Orleans area. With: Recent models indicate that there is a 1 in 5 chance that this storm will cause significant storm surge in the New Orleans area.	Pighin et al. (2011)
Include information about the reference class when using probability information.	Replace: There is a low chance of tornadoes in the Oklahoma City metro area tomorrow afternoon and evening. With: There is a low (2%) chance of tornadoes in the Oklahoma City metro area tomorrow afternoon and evening; on 1 in 50 days like today, there will be a tornado within 25 miles of your location.	Gigerenzer et. al. (2005); Juanchich & Sirota (2018)
Be aware of directionality when using probability information; positive frames can promote comprehension by encouraging people to focus on the events that are most likely to happen whereas negative frames can promote risk aversion by encouraging people to focus on the possibility of negative events, even if they are unlikely.	Positive frame: There is a high (90%) chance of sunny skies today. Negative frame: There is a low (10%) chance of freezing drizzle today.	Honda & Yamagishi (2017); Teigen & Brun (2003)
When possible, include probability information in forecast visualizations.	Supplement: Maps showing deterministic warning boxes/polygons.	Gerst et al. (2020)

	With: Maps showing probability information; for example: probability grids (FACETs); Potential Storm Surge Flooding map (NHC).	
Use visualizations to increase comprehension of probability information; icon arrays and ensemble plots can be especially effective when teaching people how to interpret probability information. Note: this might not be appropriate immediately before or during high impact events.	Supplement: Models indicate that there is an 80% to 90% chance that hurricane force winds will affect Miami, FL in the next 5 days. With: Probabilistic wind speed graphic with an icon array that deconstructs the probability information in the product.	Toet et al. (2019); Garcia-Retamero & Cokely (2014)
Pay attention to the audience when using probability information; most people are able to grasp basic probability information, but depth of comprehension depends on numeracy and experience.		Wernstedt et al. (2018); Dieckmann et al. (2009)

5. Interactive Platform

As we explain above, it is not possible to list and meaningfully summarize all of the findings from this review within the confines of a single report. Moreover, this is a living systematic review that will grow as scholars continue to study the inclusion of probability information in risk communication messages. Because of this, we believe that the results of this review are best shown in an interactive format on a platform that we update as new research becomes available. The *ProbCom* platform (short for "Probability Communication") serves this purpose. A snapshot of the interactive platform is shown in Fig. 5 and it is accessible here: https://crcm.shinyapps.io/probcom.

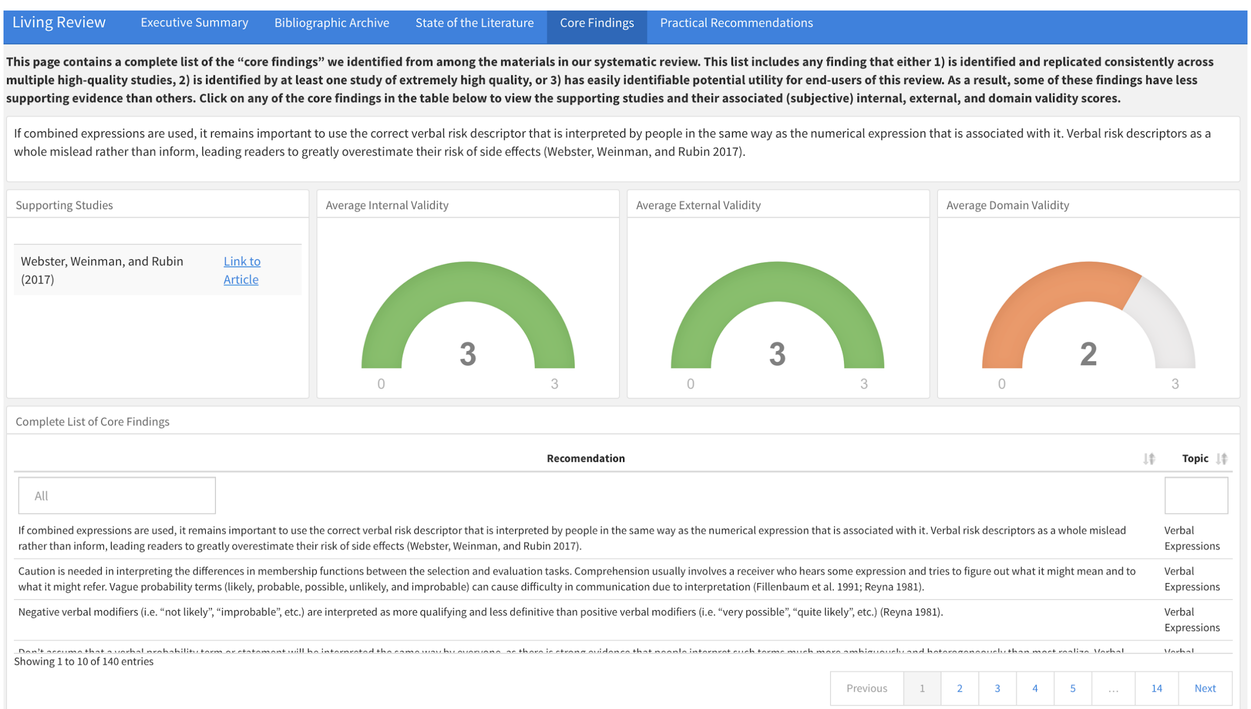

Fig. 5: Snapshot of the ProbCom Interactive Web Platform

ProbCom will evolve as new research and feedback become available, but for now it includes 5 primary tabs: an *executive summary* tab that describes the project, a *bibliographic archive* tab that lists every study in the review along with an assessment of the study's validity and online link to the study, a *state of the literature* tab that numerically and graphically summarizes data about the primary and secondary topics in the literature, a *core findings* tab that details all of the common findings we see in the literature along with assessments of mean validity, and a *recommendations* tab that highlights key recommendations for forecasters that are working to include probability information in risk messages.

6. Conclusions and Next Steps

Probabilistic forecast information is rapidly proliferating and many forecasters are struggling to figure out the best or most effective way to include probability information in risk communication. The living systematic review we introduce in this report is meant to assist in this struggle by documenting the evidence base for different communication practices. The review is ongoing and it is difficult to meaningfully synthesize everything we are learning, but a few overarching points warrant emphasis. First and foremost, this is a considerable body of research that addresses many of the questions that forecasters are grappling with. Next, in nearly all domains, the research strongly suggests that (1) average people are able to make sense of and use probability information as long as consideration is given to information presentation and (2) assuming appropriate presentation, probability information generally improves decision quality. These findings obviously beg questions about how to present probability information. We hope the recommendations and examples we provide will assist in answering some of these questions, but recognize that we are not able to speak to every question forecasters might have about probability communication. In some cases this is because we cannot summarize all of the literature in a single report. In these cases, we hope that forecasters and others will use ProbCom to find potentially relevant answers to specific questions. In other cases, we cannot speak to questions because there are relatively few (if any) quality studies that are applicable to the weather domain. These areas require future research and while there are many different ways to prioritize research needs, the findings in this review suggest that top priorities ought to include (1) more research on the use of specific words and phrases (perhaps in combination with numbers and ranges) in probability communication and (2) more research on the use of basic visualizations such as icon arrays to assist in communicating probability information to portions of the population that struggle with words and numbers; and (3) more research on how members of the public integrate (and possibly conflate) probability and severity information when judging and making decisions about low probability, high impact events.

Finally, before we close, it is important to mention a few limitations of this review and notes of caution. The first and most obvious limitation is that we necessarily simplify complex arguments and findings when aggregating and summarizing results across the studies. While necessary for synthesis, this process results in an important loss of nuance that may have implications for the recommendations we derive. We therefore strongly encourage forecasters who wish to implement the recommendations we provide to carefully review relevant studies and take note of nuance when developing implementation plans. This is one reason we provide the exemplary studies in Table 1. We believe that these studies warrant especially close attention. Next, we caution that research on *probability communication using visualizations* and *probability communication to heterogeneous populations* is particularly difficult to synthesize because most studies in these areas test specific communication strategies for specific risks, purposes, and audiences. It is therefore extremely difficult to identify systematic similarities and abstract common findings from the studies. We do the best we can, but we recognize that the common (core) findings and recommendations we make from these areas of the literature are a bit vague and less prescriptive than others. With time and feedback, we hope to clarify some of these vagaries. In the meantime, we suggest that forecasters use ProbCom to review the studies that look most interesting and relevant to the problems they are trying to solve -- there are many high quality studies on these topics and we are confident that some of them will help forecasters address some of the communication challenges they are facing.

7. References

Abraham, S., R. Bartlett, M. Standage, A. Black, A. Charlton-Perez, and R. McCloy, 2015: Do location-specific forecasts pose a new challenge for communicating uncertainty?: Communicating uncertainty and location-specific forecasts. *Met. Apps*, **22**, 554–562, https://doi.org/10.1002/met.1487.

Ancker, J. S., E. U. Weber, and R. Kukafka, 2011: Effects of Game-Like Interactive Graphics on Risk Perceptions and Decisions. *Med Decis Making*, **31**, 130–142, https://doi.org/10.1177/0272989X10364847.

Armstrong, K., G. FitzGerald, J. S. Schwartz, and P. A. Ubel, 2001: Using survival curve comparisons to inform patient decision making: Can a practice exercise improve understanding? *J Gen Intern Med*, **16**, 482–485, https://doi.org/10.1046/j.1525-1497.2001.016007482.x.

——, J. S. Schwartz, G. Fitzgerald, M. Putt, and P. A. Ubel, 2002: Effect of Framing as Gain versus Loss on Understanding and Hypothetical Treatment Choices: Survival and Mortality Curves. *Med Decis Making*, **22**, 76–83, https://doi.org/10.1177/0272989X0202200108.

Ash, K. D., R. L. Schumann, and G. C. Bowser, 2014: Tornado Warning Trade-Offs: Evaluating Choices for Visually Communicating Risk. *Weather, Climate, and Society*, **6**, 104–118, https://doi.org/10.1175/WCAS-D-13-00021.1.

Bagchi, R., and E. C. Ince, 2016: Is A 70% Forecast more Accurate than a 30% Forecast? How Level of a Forecast Affects Inferences about Forecasts and Forecasters. *Journal of Marketing Research*, **53**, 31–45, https://doi.org/10.1509/jmr.12.0526.

Baker, E. J., 1995: Public Response to Hurricane Probability Forecasts. *The Professional Geographer*, **47**, 137–147, https://doi.org/10.1111/j.0033-0124.1995.00137.x.

Ballard, T., and S. Lewandowsky, 2015: When, not if: the inescapability of an uncertain climate future. *Phil. Trans. R. Soc. A.*, **373**, 20140464, https://doi.org/10.1098/rsta.2014.0464.

Banks, S. M., P. Salovey, S. Greener, A. J. Rothman, A. Moyer, J. Beauvais, and E. Epel, 1995: The effects of message framing on mammography utilization. *Health Psychology*, **14**, 178–184, https://doi.org/10.1037/0278-6133.14.2.178.

Barnes, A. J., Y. Hanoch, T. Miron-Shatz, and E. M. Ozanne, 2016: Tailoring risk communication to improve comprehension: Do patient preferences help or hurt? *Health Psychology*, **35**, 1007–1016, https://doi.org/10.1037/hea0000367.

Bell, H., and G. Tobin, 2007: Efficient and effective? The 100-year flood in the communication and perception of flood risk. *Environmental Hazards*, **7**, 302–311, https://doi.org/10.1016/j.envhaz.2007.08.004.

Bisantz, A. M., S. S. Marsiglio, and J. Munch, 2005: Displaying Uncertainty: Investigating the Effects of Display Format and Specificity. *Hum Factors*, **47**, 777–796, https://doi.org/10.1518/001872005775570916.

Bolton, G. E., and E. Katok, 2018: Cry Wolf or Equivocate? Credible Forecast Guidance in a Cost-Loss Game. *Management Science*, **64**, 1440–1457, https://doi.org/10.1287/mnsc.2016.2645.

Bonnefon, J.-F., and G. Villejoubert, 2006: Tactful or Doubtful?: Expectations of Politeness Explain the Severity Bias in the Interpretation of Probability Phrases. *Psychol Sci*, **17**, 747–751, https://doi.org/10.1111/j.1467-9280.2006.01776.x.

Bramwell, R., H. West, and P. Salmon, 2006: Health professionals' and service users' interpretation of screening test results: experimental study. *BMJ*, **333**, 284, https://doi.org/10.1136/bmj.38884.663102.AE.

Budescu, D. V., T. M. Karelitz, and T. S. Wallsten, 2003: Predicting the directionality of probability words from their membership functions. *J. Behav. Decis. Making*, **16**, 159–180, https://doi.org/10.1002/bdm.440.

——, S. Broomell, and H.-H. Por, 2009: Improving Communication of Uncertainty in the Reports of the Intergovernmental Panel on Climate Change. *Psychol Sci*, **20**, 299–308, https://doi.org/10.1111/j.1467-9280.2009.02284.x.

——, H.-H. Por, and S. B. Broomell, 2012: Effective communication of uncertainty in the IPCC reports. *Climatic Change*, **113**, 181–200, https://doi.org/10.1007/s10584-011-0330-3.

——, ——, ——, and M. Smithson, 2014: The interpretation of IPCC probabilistic statements around the world. *Nature Clim Change*, **4**, 508–512, https://doi.org/10.1038/nclimate2194.

Carey, M., A. Herrmann, A. Hall, E. Mansfield, and K. Fakes, 2018: Exploring health literacy and preferences for risk communication among medical oncology patients. *PLoS ONE*, **13**, e0203988, https://doi.org/10.1371/journal.pone.0203988.

Cheong, L., C. Kinkeldey, I. Burfurd, S. Bleisch, and M. Duckham, 2020: Evaluating the impact of visualization of risk upon emergency route-planning. *International Journal of Geographical Information Science*, **34**, 1022–1050, https://doi.org/10.1080/13658816.2019.1701677.

Chua, H. F., J. F. Yates, and P. Shah, 2006: Risk avoidance: Graphs versus numbers. *Memory & Cognition*, **34**, 399–410, https://doi.org/10.3758/BF03193417.

Connelly, N. A., and B. A. Knuth, 1998: Evaluating Risk Communication: Examining Target Audience Perceptions About Four Presentation Formats for Fish Consumption Health Advisory Information. *Risk Analysis*, **18**, 649–659, https://doi.org/10.1111/j.1539-6924.1998.tb00377.x.

Correll, M., and M. Gleicher, 2014: Error Bars Considered Harmful: Exploring Alternate Encodings for Mean and Error. *IEEE Trans. Visual. Comput. Graphics*, **20**, 2142–2151, https://doi.org/10.1109/TVCG.2014.2346298.

Cuite, C. L., N. D. Weinstein, K. Emmons, and G. Colditz, 2008: A Test of Numeric Formats for Communicating Risk Probabilities. *Med Decis Making*, **28**, 377–384, https://doi.org/10.1177/0272989X08315246.

Denes-Raj, V., and S. Epstein, 1994: Conflict between intuitive and rational processing: When people behave against their better judgment. *Journal of Personality and Social Psychology*, **66**, 819–829, https://doi.org/10.1037/0022-3514.66.5.819.

——, ——, and J. Cole, 1995: The Generality of the Ratio-Bias Phenomenon. *Pers Soc Psychol Bull*, **21**, 1083–1092, https://doi.org/10.1177/01461672952110009.

Dieckmann, N. F., P. Slovic, and E. M. Peters, 2009: The Use of Narrative Evidence and Explicit Likelihood by Decisionmakers Varying in Numeracy. *Risk Analysis*, **29**, 1473–1488, https://doi.org/10.1111/j.1539-6924.2009.01279.x.

——, E. Peters, and R. Gregory, 2015: At Home on the Range? Lay Interpretations of Numerical Uncertainty Ranges: At Home on the Range? *Risk Analysis*, **35**, 1281–1295, https://doi.org/10.1111/risa.12358.

——, B. B. Johnson, R. Gregory, M. Mayorga, P. K. J. Han, and P. Slovic, 2017: Public perceptions of expert disagreement: Bias and incompetence or a complex and random world? *Public Underst Sci*, **26**, 325–338, https://doi.org/10.1177/0963662515603271.

Dorval, M., K. Bouchard, J. Chiquette, G. Glendon, C. M. Maugard, W. Dubuisson, S. Panchal, and J. Simard, 2013: A focus group study on breast cancer risk presentation: one format does not fit all. *Eur J Hum Genet*, **21**, 719–724, https://doi.org/10.1038/ejhg.2012.248.

Dowen, F., K. Sidhu, E. Broadbent, and H. Pilmore, 2017: Communicating projected survival with treatments for chronic kidney disease: patient comprehension and perspectives on visual aids. *BMC Med Inform Decis Mak*, **17**, 137, https://doi.org/10.1186/s12911-017-0536-z.

Doyle, E. E. H., J. McClure, D. M. Johnston, and D. Paton, 2014: Communicating likelihoods and probabilities in forecasts of volcanic eruptions. *Journal of Volcanology and Geothermal Research*, **272**, 1–15, https://doi.org/10.1016/j.jvolgeores.2013.12.006.

Durbach, I. N., and T. J. Stewart, 2011: An experimental study of the effect of uncertainty representation on decision making. *European Journal of Operational Research*, **214**, 380–392, https://doi.org/10.1016/j.ejor.2011.04.021.

Etnel, J. R. G., J. M. de Groot, M. El Jabri, A. Mesch, N. A. Nobel, A. J. J. C. Bogers, and J. J. M. Takkenberg, 2020: Do risk visualizations improve the understanding of numerical risks? A randomized, investigator-blinded general population survey. *International Journal of Medical Informatics*, **135**, 104005, https://doi.org/10.1016/j.ijmedinf.2019.104005.

Evans, J. St. B. T., S. J. Handley, N. Perham, D. E. Over, and V. A. Thompson, 2000: Frequency versus probability formats in statistical word problems. *Cognition*, **77**, 197–213, https://doi.org/10.1016/S0010-0277(00)00098-6.

Fillenbaum, S., T. S. Wallsten, B. L. Cohen, and J. A. Cox, 1991: Some Effects of Vocabulary and Communication Task on the Understanding and Use of Vague Probability Expressions. *The American Journal of Psychology*, **104**, 35, https://doi.org/10.2307/1422850.

Fischer, K., and H. Jungerman, 1996: Rarely Occurring Headaches and Rarely Occurring Blindness: Is Rarely=Rarely? The Meaning of Verbal Frequentistic Labels in Specific Medical Contexts. *Behavioral Decision Making*, **9**, 153–172, https://doi.org/10.1002/(SICI)1099-0771(199609)9:3%3C153::AID-BDM222%3E3.0.CO;2-W.

Flugstad, A. R., and P. D. Windschitl, 2003: The influence of reasons on interpretations of probability forecasts. *J. Behav. Decis. Making*, **16**, 107–126, https://doi.org/10.1002/bdm.437.

Fortin, J. M., L. K. Hirota, B. E. Bond, A. M. O'Connor, and N. F. Col, 2001: Identifying patient preferences for communicating risk estimates: A descriptive pilot study. *BMC Med Inform Decis Mak*, **1**, 2, https://doi.org/10.1186/1472-6947-1-2.

Franic, D. M., and D. S. Pathak, 2000: Communicating The Frequency of Adverse Drug Reactions to Female Patients. *Drug Information J*, **34**, 251–272, https://doi.org/10.1177/009286150003400134.

Fuller, R., 2001: Risk communication and older people--understanding of probability and risk information by medical inpatients aged 75 years and older. *Age and Ageing*, **30**, 473–476, https://doi.org/10.1093/ageing/30.6.473.

Fuller, R., 2002: How informed is consent? Understanding of pictorial and verbal probability information by medical inpatients. *Postgraduate Medical Journal*, **78**, 543–544, https://doi.org/10.1136/pmj.78.923.543.

Galesic, M., R. Garcia-Retamero, and G. Gigerenzer, 2009: Using icon arrays to communicate medical risks: Overcoming low numeracy. *Health Psychology*, **28**, 210–216, https://doi.org/10.1037/a0014474.

Garcia-Retamero, R., and M. Galesic, 2009: Communicating Treatment Risk Reduction to People With Low Numeracy Skills: A Cross-Cultural Comparison. *Am J Public Health*, **99**, 2196–2202, https://doi.org/10.2105/AJPH.2009.160234.

——, and ——, 2010: Who proficts from visual aids: Overcoming challenges in people's understanding of risks. *Social Science & Medicine*, **70**, 1019–1025, https://doi.org/10.1016/j.socscimed.2009.11.031.

——, and E. T. Cokely, 2011: Effective communication of risks to young adults: Using message framing and visual aids to increase condom use and STD screening. *Journal of Experimental Psychology: Applied*, **17**, 270–287, https://doi.org/10.1037/a0023677.

——, and M. K. Dhami, 2013: On avoiding framing effects in experienced decision makers. *Quarterly Journal of Experimental Psychology*, **66**, 829–842, https://doi.org/10.1080/17470218.2012.727836.

——, and E. T. Cokely, 2014: The Influence of Skills, Message Frame, and Visual Aids on Prevention of Sexually Transmitted Diseases: Individual Differences and Visual Aids. *J. Behav. Dec. Making*, **27**, 179–189, https://doi.org/10.1002/bdm.1797.

——, M. Galesic, and G. Gigerenzer, 2010: Do Icon Arrays Help Reduce Denominator Neglect? *Med Decis Making*, **30**, 672–684, https://doi.org/10.1177/0272989X10369000.

Gardner, P. H., B. McMillan, D. K. Raynor, E. Woolf, and P. Knapp, 2011: The effect of numeracy on the comprehension of information about medicines in users of a patient information website. *Patient Education and Counseling*, **83**, 398–403, https://doi.org/10.1016/j.pec.2011.05.006.

Gerst, M. D., and Coauthors, 2020: Using Visualization Science to Improve Expert and Public Understanding of Probabilistic Temperature and Precipitation Outlooks. *Weather, Climate, and Society*, **12**, 117–133, https://doi.org/10.1175/WCAS-D-18-0094.1.

Gibson, J. M., A. Rowe, E. R. Stone, and W. Bruine de Bruin, 2013: Communicating Quantitative Information About Unexploded Ordnance Risks to the Public. *Environ. Sci. Technol.*, **47**, 4004–4013, https://doi.org/10.1021/es305254j.

Gigerenzer, G., R. Hertwig, E. van den Broek, B. Fasolo, and K. V. Katsikopoulos, 2005: "A 30% Chance of Rain Tomorrow": How Does the Public Understand Probabilistic Weather Forecasts? *Risk Analysis*, **25**, 623–629, https://doi.org/10.1111/j.1539-6924.2005.00608.x.

Grimes, D., 1999: Patients' understanding of medical risks: implications for genetic counseling. *Obstetrics & Gynecology*, **93**, 910–914, https://doi.org/10.1016/S0029-7844(98)00567-5.

Grimm, L., E. Jessat, T. Kotovnykova, L. Otten, K. Pauli, and H. Strothmann, n.d.: A Heart for Women: How to Design Information Material on Women's Cardiac Health.

Grounds, M. A., and S. L. Joslyn, 2018: Communicating weather forecast uncertainty: Do individual differences matter? *Journal of Experimental Psychology: Applied*, **24**, 18–33, https://doi.org/10.1037/xap0000165.

——, S. Joslyn, and K. Otsuka, 2017: Probabilistic Interval Forecasts: An Individual Differences Approach to Understanding Forecast Communication. *Advances in Meteorology*, **2017**, 1–18, https://doi.org/10.1155/2017/3932565.

——, J. E. LeClerc, and S. Joslyn, 2018: Expressing Flood Likelihood: Return Period versus Probability. *Weather, Climate, and Society*, **10**, 5–17, https://doi.org/10.1175/WCAS-D-16-0107.1.

Han, P. K. J., W. M. P. Klein, T. Lehman, B. Killam, H. Massett, and A. N. Freedman, 2011: Communication of Uncertainty Regarding Individualized Cancer Risk Estimates: Effects and Influential Factors. *Med Decis Making*, **31**, 354–366, https://doi.org/10.1177/0272989X10371830.

Harris, A. J. L., and A. Corner, 2011: Communicating environmental risks: Clarifying the severity effect in interpretations of verbal probability expressions. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **37**, 1571–1578, https://doi.org/10.1037/a0024195.

Harris, A. J. L., A. Corner, and U. Hahn, 2009: Estimating the probability of negative events. *Cognition*, **110**, 51–64, https://doi.org/10.1016/j.cognition.2008.10.006.

Harris, A. J. L., A. Corner, J. Xu, and X. Du, 2013: Lost in translation? Interpretations of the probability phrases used by the Intergovernmental Panel on Climate Change in China and the UK. *Climatic Change*, **121**, 415–425, https://doi.org/10.1007/s10584-013-0975-1.

——, H.-H. Por, and S. B. Broomell, 2017: Anchoring climate change communications. *Climatic Change*, **140**, 387–398, https://doi.org/10.1007/s10584-016-1859-y.

Hawley, S. T., B. Zikmund-Fisher, P. Ubel, A. Jancovic, T. Lucas, and A. Fagerlin, 2008: The impact of the format of graphical presentation on health-related knowledge and treatment choices. *Patient Education and Counseling*, **73**, 448–455, https://doi.org/10.1016/j.pec.2008.07.023.

Hendrickx, L., C. Vlek, and H. Oppewal, 1989: Relative importance of scenario information and frequency information in the judgment of risk. *Acta Psychologica*, **72**, 41–63, https://doi.org/10.1016/0001-6918(89)90050-4.

Hess, R., V. H. M. Visschers, M. Siegrist, and C. Keller, 2011: How do people perceive graphical risk communication? The role of subjective numeracy. *Journal of Risk Research*, **14**, 47–61, https://doi.org/10.1080/13669877.2010.488745.

Hill, S., J. Spink, D. Cadilhac, A. Edwards, C. Kaufman, S. Rogers, R. Ryan, and A. Tonkin, 2010: Absolute risk representation in cardiovascular disease prevention: comprehension and preferences of health care consumers and general practitioners involved in a focus group study. *BMC Public Health*, **10**, 108, https://doi.org/10.1186/1471-2458-10-108.

Hogarth, R. M., and E. Soyer, 2015: Communicating forecasts: The simplicity of simulated experience. *Journal of Business Research*, **68**, 1800–1809, https://doi.org/10.1016/j.jbusres.2015.03.039.

Hohle, S. M., and K. H. Teigen, 2015: Forecasting forecasts: The trend effect. *Judgment and Decision Making*, **10**, 416–428.

——, and ——, 2018: More than 50% or Less than 70% Chance: Pragmatic Implications of Single-Bound Probability Estimates: Single-Bound Probability Estimates. *J. Behav. Dec. Making*, **31**, 138–150, https://doi.org/10.1002/bdm.2052.

Honda, H., and K. Yamagishi, 2006: Directional Verbal Probabilities: Inconsistencies Between Preferential Judgments and Numerical Meanings. *Experimental Psychology*, **53**, 161–170, https://doi.org/10.1027/1618-3169.53.3.161.

——, and ——, 2009: Perceived certainty based on verbal probability phrases: Effect of directionality and its dependence on method. *Japanese Psychological Research*, **51**, 266–273, https://doi.org/10.1111/j.1468-5884.2009.00409.x.

——, and ——, 2017: Communicative Functions of Directional Verbal Probabilities: Speaker's Choice, Listener's Inference, and Reference Points. *Quarterly Journal of Experimental Psychology*, **70**, 2141–2158, https://doi.org/10.1080/17470218.2016.1225779.

Howe, L. C., B. MacInnis, J. A. Krosnick, E. M. Markowitz, and R. Socolow, 2019: Acknowledging uncertainty impacts public acceptance of climate scientists' predictions. *Nat. Clim. Chang.*, **9**, 863–867, https://doi.org/10.1038/s41558-019-0587-5.

Ilic, D., K. Murphy, and S. Green, 2012: Risk Communication and Prostate Cancer: Identifying Which Summary Statistics Are Best Understood by Men. *Am J Mens Health*, **6**, 497–504, https://doi.org/10.1177/1557988312453616.

Joslyn, S., and S. Savelli, 2010: Communicating forecast uncertainty: public perception of weather forecast uncertainty. *Met. Apps*, **17**, 180–195, https://doi.org/10.1002/met.190.

——, and R. Demnitz, 2019: Communicating Climate Change: Probabilistic Expressions and Concrete Events. *Weather, Climate, and Society*, **11**, 651–664, https://doi.org/10.1175/WCAS-D-18-0126.1.

——, K. Pak, D. Jones, J. Pyles, and E. Hunt, 2007: The Effect of Probabilistic Information on Threshold Forecasts. *Weather and Forecasting*, **22**, 804–812, https://doi.org/10.1175/WAF1020.1.

——, L. Nemec, and S. Savelli, 2013: The Benefits and Challenges of Predictive Interval Forecasts and Verification Graphics for End Users. *Weather, Climate, and Society*, **5**, 133–147, https://doi.org/10.1175/WCAS-D-12-00007.1.

Joslyn, S. L., and J. E. LeClerc, 2012: Uncertainty forecasts improve weather-related decisions and attenuate the effects of forecast error. *Journal of Experimental Psychology: Applied*, **18**, 126–140, https://doi.org/10.1037/a0025185.

——, and M. A. Grounds, 2015: The use of uncertainty forecasts in complex decision tasks and various weather conditions. *Journal of Experimental Psychology: Applied*, **21**, 407–417, https://doi.org/10.1037/xap0000064.

——, and J. E. LeClerc, 2016: Climate Projections and Uncertainty Communication. *Top Cogn Sci*, **8**, 222–241, https://doi.org/10.1111/tops.12177.

——, L. Nadav-Greenberg, M. U. Taing, and R. M. Nichols, 2009: The effects of wording on the understanding and use of uncertainty information in a threshold forecasting decision. *Appl. Cognit. Psychol.*, **23**, 55–72, https://doi.org/10.1002/acp.1449.

Juanchich, M., and M. Sirota, 2016: How to improve people's interpretation of probabilities of precipitation. *Journal of Risk Research*, **19**, 388–404, https://doi.org/10.1080/13669877.2014.983945.

——, and ——, 2017: How much will the sea level rise? Outcome selection and subjective probability in climate change predictions. *Journal of Experimental Psychology: Applied*, **23**, 386–402, https://doi.org/10.1037/xap0000137.

——, and ——, 2019: Not as gloomy as we thought: reassessing how the public understands probability of precipitation forecasts. *Journal of Cognitive Psychology*, **31**, 116–129, https://doi.org/10.1080/20445911.2018.1553884.

——, ——, T. M. Karelitz, and G. Villejoubert, 2013: Can membership-functions capture the directionality of verbal probabilities? *Thinking & Reasoning*, **19**, 231–247, https://doi.org/10.1080/13546783.2013.772538.

Keller, C., and M. Siegrist, 2009: Effect of Risk Communication Formats on Risk Perception Depending on Numeracy. *Med Decis Making*, **29**, 483–490, https://doi.org/10.1177/0272989X09333122.

——, ——, and H. Gutscher, 2006: The Role of the Affect and Availability Heuristics in Risk Communication. *Risk Analysis*, **26**, 631–639, https://doi.org/10.1111/j.1539-6924.2006.00773.x.

——, ——, and V. Visschers, 2009: Effect of Risk Ladder Format on Risk Perception in High- and Low-Numerate Individuals. *Risk Analysis*, **29**, 1255–1264, https://doi.org/10.1111/j.1539-6924.2009.01261.x.

Keren, G., and L. E. M. Gerritsen, 1999: On the robustness and possible accounts of ambiguity aversion. *Acta Psychologica*, **103**, 149–172, https://doi.org/10.1016/S0001-6918(99)00034-7.

Kim, I.-G., J.-Y. Kim, B.-J. Kim, and K.-K. Lee, 2014: The collective value of weather probabilistic forecasts according to public threshold distribution patterns: Collective value of weather forecasts. *Met. Apps*, **21**, 795–802, https://doi.org/10.1002/met.1424.

Klockow-McClain, K. E., R. A. McPherson, and R. P. Thomas, 2020: Cartographic Design for Improved Decision Making: Trade-Offs in Uncertainty Visualization for Tornado Threats. *Annals of the American Association of Geographers*, **110**, 314–333, https://doi.org/10.1080/24694452.2019.1602467.

Knapp, P., P. H. Gardner, N. Carrigan, D. K. Raynor, and E. Woolf, 2009: Perceived risk of medicine side effects in users of a patient information website: A study of the use of verbal descriptors, percentages and natural frequencies. *British Journal of Health Psychology*, **14**, 579–594, https://doi.org/10.1348/135910708X375344.

Knapp, P., P. H. Gardner, D. K. Raynor, E. Woolf, and B. McMillan, 2010: Perceived risk of tamoxifen side effects: A study of the use of absolute frequencies or frequency bands, with or without verbal descriptors. *Patient Education and Counseling*, **79**, 267–271, https://doi.org/10.1016/j.pec.2009.10.002.

——, P. H. Gardner, and E. Woolf, 2016: Combined verbal and numerical expressions increase perceived risk of medicine side-effects: a randomized controlled trial of EMA recommendations. *Health Expect*, **19**, 264–274, https://doi.org/10.1111/hex.12344.

Koehler, J. J., 2001: When are people persuaded by DNA match statistics? *Law and Human Behavior*, **25**, 493–513, https://doi.org/10.1023/A:1012892815916.

Kong, A., G. O. Barnett, F. Mosteller, and C. Youtz, 1986: How Medical Professionals Evaluate Expressions of Probability. *N Engl J Med*, **315**, 740–744, https://doi.org/10.1056/NEJM198609183151206.

Kox, T., L. Gerhold, and U. Ulbrich, 2015: Perception and use of uncertainty in severe weather warnings by emergency services in Germany. *Atmospheric Research*, **158–159**, 292–301, https://doi.org/10.1016/j.atmosres.2014.02.024.

Kreye, M. E., Y. M. Goh, L. B. Newnes, and P. Goodwin, 2012: Approaches to displaying information to assist decisions under uncertainty. *Omega*, **40**, 682–692, https://doi.org/10.1016/j.omega.2011.05.010.

Kuhn, K. M., 1997: Communicating Uncertainty: Framing Effects on Responses to Vague Probabilities. *Organizational Behavior and Human Decision Processes*, **71**, 55–83, https://doi.org/10.1006/obhd.1997.2715.

Kupor, D., and K. Laurin, 2020: Probable Cause: The Influence of Prior Probabilities on Forecasts and Perceptions of Magnitude. *Journal of Consumer Research*, **46**, 833–852, https://doi.org/10.1093/jcr/ucz025.

LeClerc, J., and S. Joslyn, 2012: Odds Ratio Forecasts Increase Precautionary Action for Extreme Weather Events. *Weather, Climate, and Society*, **4**, 263–270, https://doi.org/10.1175/WCAS-D-12-00013.1.

Leitner, M., and B. P. Buttenfield, 2000: Guidelines for the Display of Attribute Certainty. *Cartography and Geographic Information Science*, **27**, 3–14, https://doi.org/10.1559/152304000783548037.

Leonhardt, J. M., and L. Robin Keller, 2018: Do Pictographs Affect Probability Comprehension and Risk Perception of Multiple-Risk Communications?: Do Pictographs Affect Probability Comprehension and Risk Perception of Multiple-Risk Communications? *J Consum Aff*, **52**, 756–769, https://doi.org/10.1111/joca.12185.

Levin, I. P., M. A. Snyder, and D. P. Chapman, 1988: The Interaction of Experiential and Situational Factors and Gender in a Simulated Risky Decision-Making Task. *The Journal of Psychology*, **122**, 173–181, https://doi.org/10.1080/00223980.1988.9712703.

Løhre, E., 2018: Stronger, sooner, and more certain climate change: A link between certainty and outcome strength in revised forecasts. *Quarterly Journal of Experimental Psychology*, **71**, 2531–2547, https://doi.org/10.1177/1747021817746062.

——, A. Sobkow, S. M. Hohle, and K. H. Teigen, 2019: Framing experts' (dis)agreements about uncertain environmental events. *J Behav Dec Making*, **32**, 564–578, https://doi.org/10.1002/bdm.2132.

Lorenz, S., S. Dessai, P. M. Forster, and J. Paavola, 2015: Tailoring the visual communication of climate projections for local adaptation practitioners in Germany and the UK. *Phil. Trans. R. Soc. A.*, **373**, 20140457, https://doi.org/10.1098/rsta.2014.0457.

Losee, J. E., K. Z. Naufel, L. Locker, and G. D. Webster, 2017: Weather Warning Uncertainty: High Severity Influences Judgment Bias. *Weather, Climate, and Society*, **9**, 441–454, https://doi.org/10.1175/WCAS-D-16-0071.1.

Maglio, S. J., and E. Polman, 2016: Revising probability estimates: Why increasing likelihood means increasing impact. *Journal of Personality and Social Psychology*, **111**, 141–158, https://doi.org/10.1037/pspa0000058.

Mandel, D. R., 2015: Accuracy of Intelligence Forecasts From the Intelligence Consumer's Perspective. *Policy Insights from the Behavioral and Brain Sciences*, **2**, 111–120, https://doi.org/10.1177/2372732215602907.

Marimo, P., T. R. Kaplan, K. Mylne, and M. Sharpe, 2015: Communication of Uncertainty in Temperature Forecasts. *Weather and Forecasting*, **30**, 5–22, https://doi.org/10.1175/WAF-D-14-00016.1.

Mazur, D. J., and D. H. Hickam, 1990: Interpretation of graphic data by patients in a general medicine clinic. *J Gen Intern Med*, **5**, 402–405, https://doi.org/10.1007/BF02599425.

McClure, J., E. E. H. Doyle, and J. M. Velluppillai, 2015: A tale of two cities: Judgments about earthquake and aftershock probabilities across time windows. *International Journal of Disaster Risk Reduction*, **14**, 15–26, https://doi.org/10.1016/j.ijdrr.2014.11.010.

McKenzie, C. R. M., and J. D. Nelson, 2003: What a speaker's choice of frame reveals: Reference points, frame selection, and framing effects. *Psychonomic Bulletin & Review*, **10**, 596–602, https://doi.org/10.3758/BF03196520.

Meyer, R., K. Broad, B. Orlove, and N. Petrovic, 2013: Dynamic Simulation as an Approach to Understanding Hurricane Risk Response: Insights from the *Stormview* Lab: Dynamic Simulation of Hurricane Risk Response. *Risk Analysis*, **33**, 1532–1552, https://doi.org/10.1111/j.1539-6924.2012.01935.x.

Miran, S. M., C. Ling, J. J. James, and L. Rothfusz, 2016: Comparing Effectiveness of Four Graphical Designs for Probabilistic Hazard Information for Tornado Threat. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, **60**, 2029–2033, https://doi.org/10.1177/1541931213601461.

——, ——, ——, A. Gerard, and L. Rothfusz, 2017: User perception and interpretation of tornado probabilistic hazard information: Comparison of four graphical designs. *Applied Ergonomics*, **65**, 277–285, https://doi.org/10.1016/j.apergo.2017.06.016.

——, ——, and L. Rothfusz, 2018: Factors influencing people's decision-making during three consecutive tornado events. *International Journal of Disaster Risk Reduction*, **28**, 150–157, https://doi.org/10.1016/j.ijdrr.2018.02.034.

Morss, R. E., J. L. Demuth, and J. K. Lazo, 2008: Communicating Uncertainty in Weather Forecasts: A Survey of the U.S. Public. *Weather and Forecasting*, **23**, 974–991, https://doi.org/10.1175/2008WAF2007088.1.

——, J. K. Lazo, and J. L. Demuth, 2010: Examining the use of weather forecasts in decision scenarios: results from a US survey with implications for uncertainty communication: USE OF WEATHER FORECASTS IN DECISION SCENARIOS. *Met. Apps*, **17**, 149–162, https://doi.org/10.1002/met.196.

——, J. L. Demuth, J. K. Lazo, K. Dickinson, H. Lazrus, and B. H. Morrow, 2016: Understanding Public Hurricane Evacuation Decisions and Responses to Forecast and Warning

Messages*. *Weather and Forecasting*, **31**, 395–417, https://doi.org/10.1175/WAF-D-15-0066.1.
Morton, T. A., A. Rabinovich, D. Marshall, and P. Bretschneider, 2011: The future that may (or may not) come: How framing changes responses to uncertainty in climate change communications. *Global Environmental Change*, **21**, 103–109, https://doi.org/10.1016/j.gloenvcha.2010.09.013.
Murphy, A. H., S. Lichtenstein, B. Fischhoff, and R. L. Winkler, 1980: Misinterpretations of Precipitation Probability Forecasts. *BAMS*, **61**, 695–701.
Nadav-Greenberg, L., and S. L. Joslyn, 2009: Uncertainty Forecasts Improve Decision Making Among Nonexperts. *Journal of Cognitive Engineering and Decision Making*, **3**, 209–227, https://doi.org/10.1518/155534309X474460.
Natter, H. M., and D. C. Berry, 2005: Effects of active information processing on the understanding of risk information. *Appl. Cognit. Psychol.*, **19**, 123–135, https://doi.org/10.1002/acp.1068.
Neace, W. P., S. Michaud, L. Bolling, K. Deer, and L. Zecevic, 2008: Frequency Formats, Probability Formats, or Problem Structure? A Test of the Nested-Sets Hypothesis in an Extensional Reasoning Task. *Judgment and Decision Making*, **3**, 140–152.
Newman, G. E., and B. J. Scholl, 2012: Bar graphs depicting averages are perceptually misinterpreted: The within-the-bar bias. *Psychon Bull Rev*, **19**, 601–607, https://doi.org/10.3758/s13423-012-0247-5.
Nurse, M. S., and W. J. Grant, 2020: I'll See It When I Believe It: Motivated Numeracy in Perceptions of Climate Change Risk. *Environmental Communication*, **14**, 184–201, https://doi.org/10.1080/17524032.2019.1618364.
Okan, Y., R. Garcia-Retamero, E. T. Cokely, and A. Maldonado, 2015: Improving risk understanding across ability levels: Encouraging active processing with dynamic icon arrays. *Journal of Experimental Psychology: Applied*, **21**, 178–194, https://doi.org/10.1037/xap0000045.
——, E. R. Stone, and W. Bruine de Bruin, 2018: Designing Graphs that Promote Both Risk Understanding and Behavior Change: Graphs Promoting Risk Understanding and Behavior Change. *Risk Analysis*, **38**, 929–946, https://doi.org/10.1111/risa.12895.
——, ——, J. Parillo, W. Bruine de Bruin, and A. M. Parker, 2020: Probability Size Matters: The Effect of Foreground-Only versus Foreground+Background Graphs on Risk Aversion Diminishes with Larger Probabilities. *Risk Analysis*, **40**, 771–788, https://doi.org/10.1111/risa.13431.
Oudhoff, J. P., and D. R. M. Timmermans, 2015: The Effect of Different Graphical and Numerical Likelihood Formats on Perception of Likelihood and Choice. *Med Decis Making*, **35**, 487–500, https://doi.org/10.1177/0272989X15576487.
Padilla, L. M., I. T. Ruginski, and S. H. Creem-Regehr, 2017: Effects of ensemble and summary displays on interpretations of geospatial uncertainty data. *Cogn. Research*, **2**, 40, https://doi.org/10.1186/s41235-017-0076-1.
Pappenberger, F., E. Stephens, J. Thielen, P. Salamon, D. Demeritt, S. J. van Andel, F. Wetterhall, and L. Alfieri, 2013: Visualizing probabilistic flood forecast information: expert preferences and perceptions of best practice in uncertainty communication:

VISUALISING PROBABILISTIC FLOOD FORECAST INFORMATION. *Hydrol. Process.*, **27**, 132–146, https://doi.org/10.1002/hyp.9253.
Patt, A. G., and D. P. Schrag, 2003: [No title found]. *Climatic Change*, **61**, 17–30, https://doi.org/10.1023/A:1026314523443.
van Pelt, S. C., M. Haasnoot, B. Arts, F. Ludwig, R. Swart, and R. Biesbroek, 2015: Communicating climate (change) uncertainties: Simulation games as boundary objects. *Environmental Science & Policy*, **45**, 41–52, https://doi.org/10.1016/j.envsci.2014.09.004.
Peters, E., D. Västfjäll, P. Slovic, C. K. Mertz, K. Mazzocco, and S. Dickert, 2006: Numeracy and Decision Making. *Psychol Sci*, **17**, 407–413, https://doi.org/10.1111/j.1467-9280.2006.01720.x.
Peters, E. M., and I. P. Levin, 2008: Dissecting the Risky-Choice Framing Effect: Numeracy as an Individual-Difference Factor in Weighting Risky and Riskless Options. *Judgment and Decision Making*, **3**, 435–448.
Piercey, M. D., 2009: Motivated reasoning and verbal vs. numerical probability assessment: Evidence from an accounting context. *Organizational Behavior and Human Decision Processes*, **108**, 330–341, https://doi.org/10.1016/j.obhdp.2008.05.004.
Pighin, S., L. Savadori, E. Barilli, L. Cremonesi, M. Ferrari, and J.-F. Bonnefon, 2011: The 1-in- *X* Effect on the Subjective Assessment of Medical Probabilities. *Med Decis Making*, **31**, 721–729, https://doi.org/10.1177/0272989X11403490.
——, ——, ——, S. Galbiati, M. Smid, M. Ferrari, and L. Cremonesi, 2015: Communicating Down syndrome risk according to maternal age: “1-in- *X* ” effect on perceived risk: “1-in- *X* ” effect on perceived risk. *Prenat Diagn*, **35**, 777–782, https://doi.org/10.1002/pd.4606.
Rabinovich, A., and T. A. Morton, 2012: Unquestioned Answers or Unanswered Questions: Beliefs About Science Guide Responses to Uncertainty in Climate Change Risk Communication: Model of Science and Response to Uncertainty. *Risk Analysis*, **32**, 992–1002, https://doi.org/10.1111/j.1539-6924.2012.01771.x.
Ramos, M. H., S. J. van Andel, and F. Pappenberger, 2013: Do probabilistic forecasts lead to better decisions? *Hydrol. Earth Syst. Sci.*, **17**, 2219–2232, https://doi.org/10.5194/hess-17-2219-2013.
Retchless, D. P., and C. A. Brewer, 2016: Guidance for representing uncertainty on global temperature change maps. *Int. J. Climatol.*, **36**, 1143–1159, https://doi.org/10.1002/joc.4408.
Reyna, V. F., 1981: The language of possibility and probability: Effects of negation on meaning. *Mem Cogn*, **9**, 642–650, https://doi.org/10.3758/BF03202359.
Riege, A. H., and K. H. Teigen, 2013: Additivity neglect in probability estimates: Effects of numeracy and response format. *Organizational Behavior and Human Decision Processes*, **121**, 41–52, https://doi.org/10.1016/j.obhdp.2012.11.004.
Rinne, L. F., and M. M. M. Mazzocco, 2013: Inferring Uncertainty from Interval Estimates: Effects of Alpha Level and Numeracy. *Judgment and Decision Making*, **8**, 330–344.
Roth, R. E., 2009: The Impact of User Expertise on Geographic Risk Assessment under Uncertain Conditions. *Cartography and Geographic Information Science*, **36**, 29–43, https://doi.org/10.1559/152304009787340160.

Roulston, M. S., and T. R. Kaplan, 2009: A laboratory-based study of understanding of uncertainty in 5-day site-specific temperature forecasts: UNDERSTANDING UNCERTAINTY IN FORECASTS. *Met. Apps*, **16**, 237–244, https://doi.org/10.1002/met.113.

——, G. E. Bolton, A. N. Kleit, and A. L. Sears-Collins, 2006: A Laboratory Study of the Benefits of Including Uncertainty Information in Weather Forecasts. *Weather and Forecasting*, **21**, 116–122, https://doi.org/10.1175/WAF887.1.

Ruginski, I. T., and Coauthors, 2016: Non-expert interpretations of hurricane forecast uncertainty visualizations. *Spatial Cognition & Computation*, **16**, 154–172, https://doi.org/10.1080/13875868.2015.1137577.

Ruiz, J. G., A. D. Andrade, R. Garcia-Retamero, R. Anam, R. Rodriguez, and J. Sharit, 2013: Communicating global cardiovascular risk: Are icon arrays better than numerical estimates in improving understanding, recall and perception of risk? *Patient Education and Counseling*, **93**, 394–402, https://doi.org/10.1016/j.pec.2013.06.026.

Sanyal, J., Song Zhang, G. Bhattacharya, P. Amburn, and R. Moorhead, 2009: A User Study to Compare Four Uncertainty Visualization Methods for 1D and 2D Datasets. *IEEE Trans. Visual. Comput. Graphics*, **15**, 1209–1218, https://doi.org/10.1109/TVCG.2009.114.

Savelli, S., and S. Joslyn, 2012: Boater Safety: Communicating Weather Forecast Information to High-Stakes End Users. *Weather, Climate, and Society*, **4**, 7–19, https://doi.org/10.1175/WCAS-D-11-00025.1.

Schapira, M. M., A. B. Nattinger, and T. L. McAuliffe, 2006: The Influence of Graphic Format on Breast Cancer Risk Communication. *Journal of Health Communication*, **11**, 569–582, https://doi.org/10.1080/10810730600829916.

Schirillo, J. A., and E. R. Stone, 2005: The Greater Ability of Graphical Versus Numerical Displays to Increase Risk Avoidance Involves a Common Mechanism. *Risk Analysis*, **25**, 555–566, https://doi.org/10.1111/j.1539-6924.2005.00624.x.

Seipel, S., and N. J. Lim, 2017: Color map design for visualization in flood risk assessment. *International Journal of Geographical Information Science*, **31**, 2286–2309, https://doi.org/10.1080/13658816.2017.1349318.

Sherman-Morris, K., and I. Del Valle-Martinez, 2017: Optimistic bias and the consistency of hurricane track forecasts. *Nat Hazards*, **88**, 1523–1543, https://doi.org/10.1007/s11069-017-2931-2.

——, K. B. Antonelli, and C. C. Williams, 2015: Measuring the Effectiveness of the Graphical Communication of Hurricane Storm Surge Threat. *Weather, Climate, and Society*, **7**, 69–82, https://doi.org/10.1175/WCAS-D-13-00073.1.

Sinayev, A., E. Peters, M. Tusler, and L. Fraenkel, 2015: Presenting Numeric Information with Percentages and Descriptive Risk Labels: A Randomized Trial. *Med Decis Making*, **35**, 937–947, https://doi.org/10.1177/0272989X15584922.

Sink, S. A., 1995: DETERMINING THE PUBLIC'S UNDERSTANDING OF PRECIPITATION FORECASTS: RESULTS OF A SURVEY. *National Weather Digest*, **19**, 7.

Smithson, M., D. V. Budescu, S. B. Broomell, and H.-H. Por, 2012: Never say "not": Impact of negative wording in probability phrases on imprecise probability judgments. *International*

Journal of Approximate Reasoning, **53**, 1262–1270, https://doi.org/10.1016/j.ijar.2012.06.019.
Stone, E. R., J. F. Yates, and A. M. Parker, 1997: Effects of numerical and graphical displays on professed risk-taking behavior. *Journal of Experimental Psychology: Applied*, **3**, 243–256, https://doi.org/10.1037/1076-898X.3.4.243.
——, W. Bruine de Bruin, A. M. Wilkins, E. M. Boker, and J. MacDonald Gibson, 2017: Designing Graphs to Communicate Risks: Understanding How the Choice of Graphical Format Influences Decision Making: Graphical Risk Communication Formats. *Risk Analysis*, **37**, 612–628, https://doi.org/10.1111/risa.12660.
Strathie, A., G. Netto, G. H. Walker, and G. Pender, 2017: How presentation format affects the interpretation of probabilistic flood risk information: Communicating probabilistic flood risk. *J Flood Risk Management*, **10**, 87–96, https://doi.org/10.1111/jfr3.12152.
Tak, S., and A. Toet, 2014: Color and Uncertainty: It is not always Black and White. *EuroVis - Short Papers*, 5 pages, https://doi.org/10.2312/EUROVISSHORT.20141157.
Taylor, B. J., M. Stevenson, and M. McDowell, 2018: Communicating risk in dementia care: Survey of health and social care professionals. *Health Soc Care Community*, **26**, https://doi.org/10.1111/hsc.12519.
Teigen, K. H., and W. Brun, 1995: Yes, but it is uncertain: Direction and communicative intention of verbal probabilistic terms. *Acta Psychologica*, **88**, 233–258, https://doi.org/10.1016/0001-6918(93)E0071-9.
——, and ——, 1999: The Directionality of Verbal Probability Expressions: Effects on Decisions, Predictions, and Probabilistic Reasoning. *Organizational Behavior and Human Decision Processes*, **80**, 155–190, https://doi.org/10.1006/obhd.1999.2857.
——, and ——, 2000: Ambiguous Probabilities: When Does P=0.3 Reflect a Possibility, and When Does It Express a Doubt? *Behavioral Decision Making*, **13**, 345–362.
——, and ——, 2003: Verbal probabilities: a question of frame? *J. Behav. Decis. Making*, **16**, 53–72, https://doi.org/10.1002/bdm.432.
——, and P. Filkuková, 2013: Can > Will: Predictions of What *Can* Happen are Extreme, but Believed to be Probable: Can > Will. *J. Behav. Dec. Making*, **26**, 68–78, https://doi.org/10.1002/bdm.761.
——, M. Juanchich, and A. H. Riege, 2013: Improbable outcomes: Infrequent or extraordinary? *Cognition*, **127**, 119–139, https://doi.org/10.1016/j.cognition.2012.12.005.
——, ——, and P. Filkuková, 2014: Verbal Probabilities: An Alternative Approach. *Quarterly Journal of Experimental Psychology*, **67**, 124–146, https://doi.org/10.1080/17470218.2013.793731.
——, P. Filkuková, and S. M. Hohle, 2018a: It can become 5 °C warmer: The extremity effect in climate forecasts. *Journal of Experimental Psychology: Applied*, **24**, 3–17, https://doi.org/10.1037/xap0000149.
——, E. Lohre, and S. M. Hohle, 2018b: The Boundary Effect: Perceived Post Hoc Accuracy of Prediction Intervals. *Judgment and Decision Making*, **13**, 309–321.
Toet, A., J. B. van Erp, E. M. Boertjes, and S. van Buuren, 2019: Graphical uncertainty representations for ensemble predictions. *Information Visualization*, **18**, 373–383, https://doi.org/10.1177/1473871618807121.

Tubau, E., J. Rodríguez-Ferreiro, I. Barberia, and À. Colomé, 2019: From reading numbers to seeing ratios: a benefit of icons for risk comprehension. *Psychological Research*, **83**, 1808–1816, https://doi.org/10.1007/s00426-018-1041-4.

Tyszka, T., and P. Sawicki, 2011: Affective and Cognitive Factors Influencing Sensitivity to Probabilistic Information: Sensitivity to Probabilistic Information. *Risk Analysis*, **31**, 1832–1845, https://doi.org/10.1111/j.1539-6924.2011.01644.x.

Ulph, F., E. Townsend, and C. Glazebrook, 2009: How should risk be communicated to children: a cross-sectional study comparing different formats of probability information. *BMC Med Inform Decis Mak*, **9**, 26, https://doi.org/10.1186/1472-6947-9-26.

Wallsten, T. S., S. Fillenbaum, and J. A. Cox, 1986: Base rate effects on the interpretations of probability and frequency expressions. *Journal of Memory and Language*, **25**, 571–587, https://doi.org/10.1016/0749-596X(86)90012-4.

Weber, E. U., and D. J. Hilton, 1990: Contextual effects in the interpretations of probability words: Perceived base rate and severity of events. *Journal of Experimental Psychology: Human Perception and Performance*, **16**, 781–789, https://doi.org/10.1037/0096-1523.16.4.781.

Wernstedt, K., P. S. Roberts, J. Arvai, and K. Redmond, 2019: How emergency managers (mis?)interpret forecasts. *Disasters*, **43**, 88–109, https://doi.org/10.1111/disa.12293.

Wilson, K. A., P. L. Heinselman, P. S. Skinner, J. J. Choate, and K. E. Klockow-McClain, 2019: Meteorologists' Interpretations of Storm-Scale Ensemble-Based Forecast Guidance. *Weather, Climate, and Society*, **11**, 337–354, https://doi.org/10.1175/WCAS-D-18-0084.1.

Windschitl, P. D., and E. U. Weber, 1999: The Interpretation of "Likely" Depends on the Context, But 70% is 70% - Right? The Influence of Associative Processes on Perceived Certainty. *SSRN Journal*, https://doi.org/10.2139/ssrn.1304758.

Wintle, B. C., H. Fraser, B. C. Wills, A. E. Nicholson, and F. Fidler, 2019: Verbal probabilities: Very likely to be somewhat more confusing than numbers. *PLoS ONE*, **14**, e0213522, https://doi.org/10.1371/journal.pone.0213522.

Witteman, H. O., A. Fuhrel-Forbis, H. C. Wijeysundera, N. Exe, M. Dickson, L. Holtzman, V. C. Kahn, and B. J. Zikmund-Fisher, 2014: Animated Randomness, Avatars, Movement, and Personalization in Risk Graphics. *J Med Internet Res*, **16**, e80, https://doi.org/10.2196/jmir.2895.

Wu, H.-C., M. K. Lindell, C. S. Prater, and C. D. Samuelson, 2014: Effects of Track and Threat Information on Judgments of Hurricane Strike Probability: Judgments of Hurricane Strike Probability. *Risk Analysis*, **34**, 1025–1039, https://doi.org/10.1111/risa.12128.

Wu, L., S. Zeng, and Y. Wu, 2018: Affect Heuristic and Format Effect in Risk Perception. *soc behav pers*, **46**, 1331–1344, https://doi.org/10.2224/sbp.6957.

Zabini, F., V. Grasso, R. Magno, F. Meneguzzo, and B. Gozzini, 2015: Communication and interpretation of regional weather forecasts: a survey of the Italian public: Survey analysis on the interpretation of weather forecasts. *Met. Apps*, **22**, 495–504, https://doi.org/10.1002/met.1480.

Zikmund-Fisher, B. J., A. Fagerlin, T. R. Roberts, H. A. Derry, and P. A. Ubel, 2008: Alternate Methods of Framing Information About Medication Side Effects: Incremental Risk Versus Total Risk of Occurrence. *Journal of Health Communication*, **13**, 107–124, https://doi.org/10.1080/10810730701854011.

Zikmund-Fisher, B. J., M. Dickson, and H. O. Witteman, 2011: Cool but Counterproductive: Interactive, Web-Based Risk Communications Can Backfire. *J Med Internet Res*, **13**, e60, https://doi.org/10.2196/jmir.1665.

Zikmund-Fisher, B. J., and Coauthors, 2014: Blocks, Ovals, or People? Icon Type Affects Risk Perceptions and Recall of Pictographs. *Med Decis Making*, **34**, 443–453, https://doi.org/10.1177/0272989X13511706.

8. Appendix A: Search Terms

ProQuest Search

S	Search Terms	# of Results
1	ti(communicat* OR perception OR inform* OR messag* OR understand*) AND PEER(yes)	679,783
2	ti(risk OR probabil* OR uncertain*) AND PEER(yes)	707,081
3	ab(weather OR climat* OR meterolog* OR "global warming" OR forecast*) AND PEER(yes)	860,375
4	(ab(experiment* OR survey* OR data* OR statistic*) OR ti(experiment* OR survey* OR data* OR statistic*)) AND PEER(yes)	13,223,159
5	S1 AND S2 AND S4	13,442
6	S1 AND S2 AND S3 AND S4	**788**
7	S1 AND S2 AND S3	1,350

Web of Science Search

S	Search Terms	# of Results
1	ti=(communicat* OR perception OR inform* OR messag* OR understand*)	733,747
2	ti=(risk OR probabil* OR uncertain*)	814,479
3	ts=(weather OR climat* OR meterolog* OR "global warming" OR forecast*)	716,340
4	ts=(experiment* OR survey* OR data* OR statistic*) OR ti=(experiment* OR survey* OR data* OR statistic*)	10,657,241
5	S1 AND S2 AND S4	**7,007**
6	S1 AND S2 AND S3 AND S4	**565**

EBSCO (Academic Search Elite) Search

S	Search Terms	# of Results
1	ti(communicat* OR perception OR inform* OR messag* OR understand*)	364,341
2	ti(risk OR probabil* OR uncertain*)	315,672
3	ab(weather OR climat* OR meterolog* OR "global warming" OR forecast*)	322,317
4	S1 AND S2 AND S3	378
5	ab(experiment* OR survey* OR data* OR statistic*) OR ti(experiment* OR survey* OR data* OR statistic*))	5,060,185
6	S1 AND S2 AND S5	**3,493**
7	S1 AND S2 AND S3 AND S5	**206**

9. Appendix B: Previous Literature Reviews

van der Bles, Anne Marthe et al. 2019. "Communicating Uncertainty about Facts, Numbers and Science." *Royal Society open science* 6(5): 181870. https://search.proquest.com/docview/2244131098?accountid=12964.

National Research Council. 2002. *Communicating Uncertainties in Weather and Climate Information: A Workshop Summary*. https://www.nap.edu/catalog/10597/communicating-uncertainties-in-weather-and-climate-information-a-workshop-summary (August 6, 2019).

National Research Council. 2007. *Completing the Forecast: Characterizing and Communicating Uncertainty for Better Decisions Using Weather and Climate Forecasts*. https://www.nap.edu/catalog/11699/completing-the-forecast-characterizing-and-communicating-uncertainty-for-better-decisions (August 6, 2019).

National Research Council. 2012. *Weather Services for the Nation: Becoming Second to None*. https://www.nap.edu/catalog/13429/weather-services-for-the-nation-becoming-second-to-none (August 6, 2019).

Fitzpatrick-Lewis, Donna, Jennifer Yost, Donna Ciliska, and Shari Krishnaratne. 2010. "Communication about Environmental Health Risks: A Systematic Review." *Environmental Health* 9(1): 67. https://doi.org/10.1186/1476-069X-9-67 (August 13, 2019).

Glik, Deborah C. 2007. "Risk Communication for Public Health Emergencies." *Annual Review of Public Health* 28: 33–54.

Gustafson, Abel, and Ronald E. Rice. 2020. "A Review of the Effects of Uncertainty in Public Science Communication." *Public Understanding of Science* 29(6): 614–33.

Hirschberg, Paul et al. 2011. "A Weather and Climate Enterprise Strategic Implementation Plan for Generating and Communicating Forecast Uncertainty Information." *Bulletin of the American Meteorological Society* 92: 1651–66.

Kellens, Wim, Teun Terpstra, and Philippe De Maeyer. 2013. "Perception and Communication of Flood Risks: A Systematic Review of Empirical Research." *Risk Analysis: An Official Publication of the Society for Risk Analysis* 33(1): 24–49.

Spiegelhalter, David. 2017. "Risk and Uncertainty Communication." In *ANNUAL REVIEW OF STATISTICS AND ITS APPLICATION, VOL 4*, Annual Review of Statistics and Its Application, ed. Fienberg, SE. , 31–60.

Spiegelhalter, David, Mike Pearson, and Ian Short. 2011. "Visualizing Uncertainty About the Future." *Science* 333(6048): 1393–1400. https://science.sciencemag.org/content/333/6048/1393 (August 6, 2019).

Visschers, Vivianne H. M., Ree M. Meertens, Wim W. F. Passchier, and Nanne N. K. De Vries. 2009. "Probability Information in Risk Communication: A Review of the Research Literature." *Risk Analysis* 29(2): 267–87. https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1539-6924.2008.01137.x (August 6, 2019).