How Visual Design of Severe Weather Outlooks Can Affect Communication and Decision-Making

MARY ANNE T. CLIVE,[a] EMMA E. H. DOYLE,[b] SALLY H. POTTER,[a] CHRIS NOBLE,[c] AND DAVID M. JOHNSTON[b]

[a] GNS Science, Te Pū Ao, Lower Hutt, New Zealand
[b] Joint Centre for Disaster Research, Massey University, Wellington, New Zealand
[c] MetService, Te Ratonga Tirorangi, Wellington, New Zealand

(Manuscript received 30 January 2023, in final form 31 July 2023, accepted 4 August 2023)

ABSTRACT: Multiday severe weather outlooks can inform planning beyond the hour-to-day windows of warnings and watches. Outlooks can be complex to visualize, as they represent large-scale weather phenomena overlapping across several days at varying levels of uncertainty. Here, we present the results of a survey ($n = 417$) that explores how visual variables affect comprehension, inferences, and intended decision-making in a hypothetical scenario with the New Zealand MetService Severe Weather Outlook. We propose that visualization of the time window, forecast area, icons, and uncertainty can influence perceptions and decision-making based on four key findings. First, composite-style outlooks that depict multiple days of weather on one map can lead to biased perceptions of the forecast. When responding to questions about a day for which participants accurately reported there was no severe weather forecast, those who viewed a composite outlook reported higher likelihoods of severe weather occurring, higher levels of concern about travel, and higher likelihoods of changing plans compared to those who viewed outlooks that showed weather for each day on a separate map, suggesting that they perceived the forecast to underrepresent the likelihood of severe weather on that day. Second, presenting uncertainty in an extrinsic way (e.g., "low") can lead to more accurate estimates of likelihood than intrinsic formats (e.g., hue variation). Third, shaded forecast areas may lead to higher levels of confidence in the forecast than outlined forecast areas. Fourth, inclusion of weather icons can improve comprehension in some conditions. The results demonstrate how visualization can affect decision-making about severe weather and support several evidence-based considerations for effective design of long-term forecasts.

SIGNIFICANCE STATEMENT: Severe weather outlook forecasts can be hard to clearly communicate because they show multiple weather patterns across multiple days and regions with varying uncertainty. The purpose of this study is to explore how visual elements of outlook design affect the way that people understand this complex content. We had three separate groups respond to the same series of questions while viewing different modified versions of the MetService Severe Weather Outlook in Aotearoa New Zealand and compared their responses. We find that the way the outlooks' time window, forecast area, icons, and uncertainty are visualized can influence how people understand outlooks and make inferences and decisions about severe weather. We discuss how these influences may impact communication and action and present several evidence-based considerations for effective outlook design.

KEYWORDS: Social science; Uncertainty; Forecasting; Communications/decision-making; Decision-making

1. Introduction

Severe weather outlooks are used to alert audiences about potential high-impact weather events several days in advance. These products, such as the National Oceanic and Atmospheric Administration (NOAA) Storm Prediction Centre (SPC) Convective Outlooks in the United States and the Meteorological Service of New Zealand, Ltd. (MetService), Severe Weather Outlook in Aotearoa New Zealand (NZ), often communicate information about multiple weather phenomena across multiple days with multiple levels of uncertainty at large regional scales. Although commonly used in planning by specialist audiences, such as emergency and risk managers, these longer-range national forecast visualizations are increasingly used and shared by public audiences on social media for activities such as planning weekend events (e.g., Ripberger et al. 2014; Ernst et al. 2021). Currently, it is unclear how the visual design and presentation of these complex outlooks influence interpretation of information and decision-making surrounding severe weather among nonspecialist audiences.

Previous studies have found that visualization of weather forecasts can have important influences on comprehension and behavioral intentions, but these effects have not been widely explored or considered in terms of the development of communication products, such as maps and graphics (Carr et al. 2016). Interpreting forecast information from a map or graphic is a complex information-processing exercise that relies on an individual's perceptual and cognitive processing of shapes, symbols, and text in the context of prior experience, existing risk beliefs, and heuristics (Dodge et al. 2011; Hegarty 2011; Severtson and Myers 2013; Patterson et al. 2014; Padilla et al. 2018). Miscommunication and barriers may arise when

Denotes content that is immediately available upon publication as open access.

Corresponding author: Mary Anne T. Clive, m.clive@gns.cri.nz

DOI: 10.1175/WCAS-D-23-0010.1

design factors impinge on this processing, for example, if the design is overly technical or confusing, has illegible text, uses nonintuitive color schemes, or has task-relevant information that is not visually salient (e.g., Broad et al. 2007; Hegarty et al. 2010; Carr et al. 2016).

These issues could have important implications for life safety, as forecast comprehension could play a role in whether or not protective actions are ultimately taken (Paton et al. 2005; Lindell and Perry 2012). The risk perceptions that evolve from cognitive processing of the forecast graphic may help prompt a behavioral response, such as taking a protective action (e.g., changing or canceling plans) or seeking more information (e.g., following weather warning and watch updates). For longer-term multiday outlooks, accurately identifying the weather processes and uncertainty information for each day is critical for developing an informed understanding of the distribution and variability in weather patterns across the outlook period. For example, this knowledge could help inform decision-making for outdoor excursions, such as multiday hikes, where decisions to visit remote outdoor locations with limited telecommunications may be made several days in advance. Between 2007 and 2017, weather was deemed to be a causal factor in 38% of all multiday hike fatalities in NZ (New Zealand Mountain Safety Council 2018). Severe weather outlooks can give early warnings to those who are planning activities in remote or risky outdoor locations. They can also help proactively inform emergency management strategies and resource planning for major weather disruptions in advance of shorter-term warnings and watches.

Previous studies have found that people generally understand that the purpose of the MetService Severe Weather Outlook in NZ is to alert about dangerous or high-impact weather that could approach within several days (Wright et al. 2010). However, comprehension of the outlook design has not yet been explored. Previous studies have demonstrated the value of empirically testing user decision-making with weather forecast designs to improve and inform products (e.g., Morss et al. 2008; Demuth et al. 2013). Potter et al. (2018) used online surveys with the NZ public to investigate the influence of severe weather impact messaging on perceptions and intended actions, and Ernst et al. (2021) suggest that user surveys could improve outlook design and reduce harm from severe weather events. Here, we use a survey to empirically test how the visual design elements of forecasts, such as the timeline layout, visualization style, and framing of uncertainty, influence the way that people understand and make decisions with regard to the NZ MetService Severe Weather Outlook graphic.

a. *MetService Severe Weather Outlook*

MetService is the official alerting authority for severe weather in NZ (NEMA 2015). MetService issues three different types of forecasts to notify about severe weather: warnings (next 24 h), watches (1–2 days out), and outlooks (3–6 days out). The MetService Severe Weather Outlook targets widespread (1000 km^2) wind, rain, and snow events in NZ that meet any of the following "severe weather" thresholds: rainfall greater than 50 mm within 6 h or 100 mm within 24 h; snowfall below 1000 m on the North Island or in the south Canterbury or Otago region of NZ or below 500 m elsewhere on the South Island of NZ, with a snow depth of 10 cm within 6 h or 25 cm within 24 h; and severe gales, with a minimum mean speed of 90 km h^{-1} or frequent gusts exceeding 110 km h^{-1}. The events are typically driven by synoptic features, such as low-pressure systems and fronts, that can be identified and predicted several days in advance. Unlike the NOAA SPC Convective Outlooks (which cover 1, 2, 3, and 4–8 days out), they do not assess population exposure or risk, and they do not include phenomena such as thunderstorms, hail, and tornadoes, which occur at a localized scale in NZ and are targeted by the MetService Thunderstorm Outlook (which covers the next 24–36 h). A 2017 survey found that the NZ public is most familiar with the MetService's Severe Weather Warning and least familiar with the longer-range Severe Weather Outlook (Perceptive 2017). Although web usage statistics indicate that Severe Weather Outlooks are accessed much less frequently than daily weather forecasts, engagement with the outlook graphics on social media can increase greatly in the lead-up to public holiday periods and long weekends. This aligns with previous studies that have found that planning weekend activities and travel is one of the most common uses of weather forecast data among public users (Lazo et al. 2009).

The MetService Severe Weather Outlook shows a national-scale map of NZ, with areas of potential large severe weather events outlined in colored hashed lines, with a connector arrow that lists the date, type of severe weather forecast (heavy rain, strong winds, or heavy snow), and associated confidence (Fig. 1; https://www.metservice.com/warnings/severe-weather-outlook). The outlook includes forecasts for 3–6 days out and is land based only, with limited geographical information (e.g., no roads or city labels) and does not include marine areas. In these outlooks, confidence is defined as one of three different likelihoods that the severe weather event will "actually happen" [low (20%), moderate (40%), or high (60%)], which is assigned through the expert judgment of experienced forecasters working on the day. The color of an outlined forecast area on the map is not explicitly linked to a meteorological process, timing, or level of confidence but is chosen from a fixed set of 15 colors at the discretion of the forecaster. Although there is no formal guidance on how to apply colors, forecasters typically use brighter colors (e.g., red or magenta) for areas of higher confidence or more extreme impacts, and they may apply similar colors to areas of the same meteorological phenomena. There are potential limitations associated with this approach, and little evidence is currently available to guide such design choices. This has in part motivated this study into how alternative visualization choices influence audience perceptions.

The outlooks show four days of forecast weather on one composite map that represents all severe weather forecast for 3, 4, 5, and 6 days out from the current day. On the MetService web site, the outlooks are accompanied by a detailed text description of the forecast and a key that defines the confidence levels and likelihoods. However, the outlook is often shared as a standalone image across other social and traditional media platforms, without the key or full accompanying text description. It is therefore important to understand how audiences engage with and understand the outlook image as an individual product. Here,

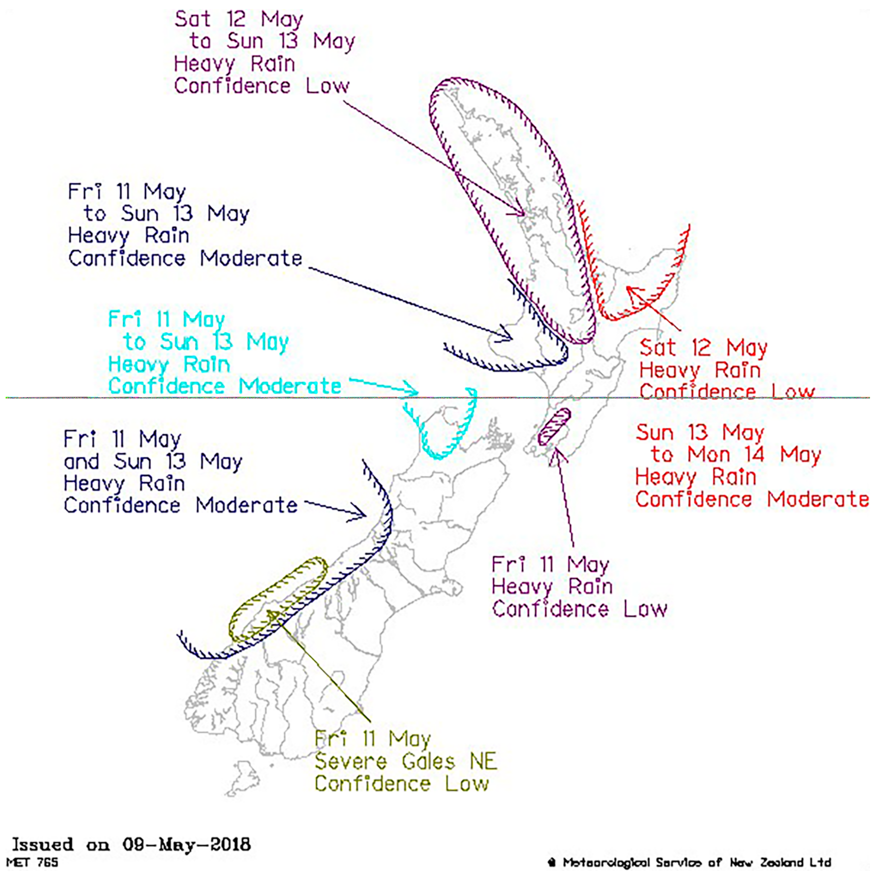

FIG. 1. MetService Severe Weather Outlook published 9 May 2018. It is the control map used to inform this study.

we explore the effect of three outlook design variables: visualization of the time frame, expressions of uncertainty, and representation of the forecast area and weather type.

b. *Time frame visualization and temporal uncertainty*

The multiday national nature of outlook forecasts requires users to extract complex visual information about different types of events over multiple days and regions in order to understand which type of severe weather is occurring on which days. To assist in making sense of such information, visual variables are processed by an individual's working memory and mentally organized into groups through a process called chunking (Miller 1956; Egan and Schwartz 1979). Chunking helps to increase the amount of information that can be encoded and stored in our fixed-capacity working memory, which can reach high cognitive loads when large amounts of complex information are being processed (Evans and Stanovich 2013; Thalmann et al. 2019). Scientific visualizations can facilitate chunking of information for readers by using visual variables and design principles, such as color and layout, to drive visual attention to relevant information (Patterson et al. 2014). Chunking can reduce complexity and clutter to help nonexpert users comprehend meteorological data more easily (Harold et al. 2016).

When considering how to present "chunked" information about temporal uncertainty, one design factor to evaluate is whether the map should be displayed in a "univariate" (designing multiple graphics for comparison side by side) or "bivariate" (designing one graphic with extrinsic and intrinsic elements to display different layers of information) manner (Kunz et al. 2011; Kubíček and Šašinka 2011; Gould et al. 2014; Kinkeldey et al. 2014; Doyle et al. 2019). Kubíček and Šašinka (2011) have identified that the latter encourages parallel processing, while the former encourages serial processing, which requires greater information to be held in the working memory for the comparison to occur. The design choices thus each have trade-offs for decision-making, with univariate maps resulting in slower but more accurate interpretations and bivariate maps resulting in quicker but less accurate interpretations.

In this study, we compare methods for perceptually chunking time in outlook forecasts to explore how the different time frame displays affect comprehension of forecast information, perceptions of temporal uncertainty, and ease of reading information from the outlook. We compare the following three conditions, illustrated in Fig. 2: 1) a composite (bivariate) outlook that shows all 4 days of the forecast on one map (the MetService Severe Weather Outlook's current approach), 2) an animated outlook (univariate) that shows a map for each day displayed

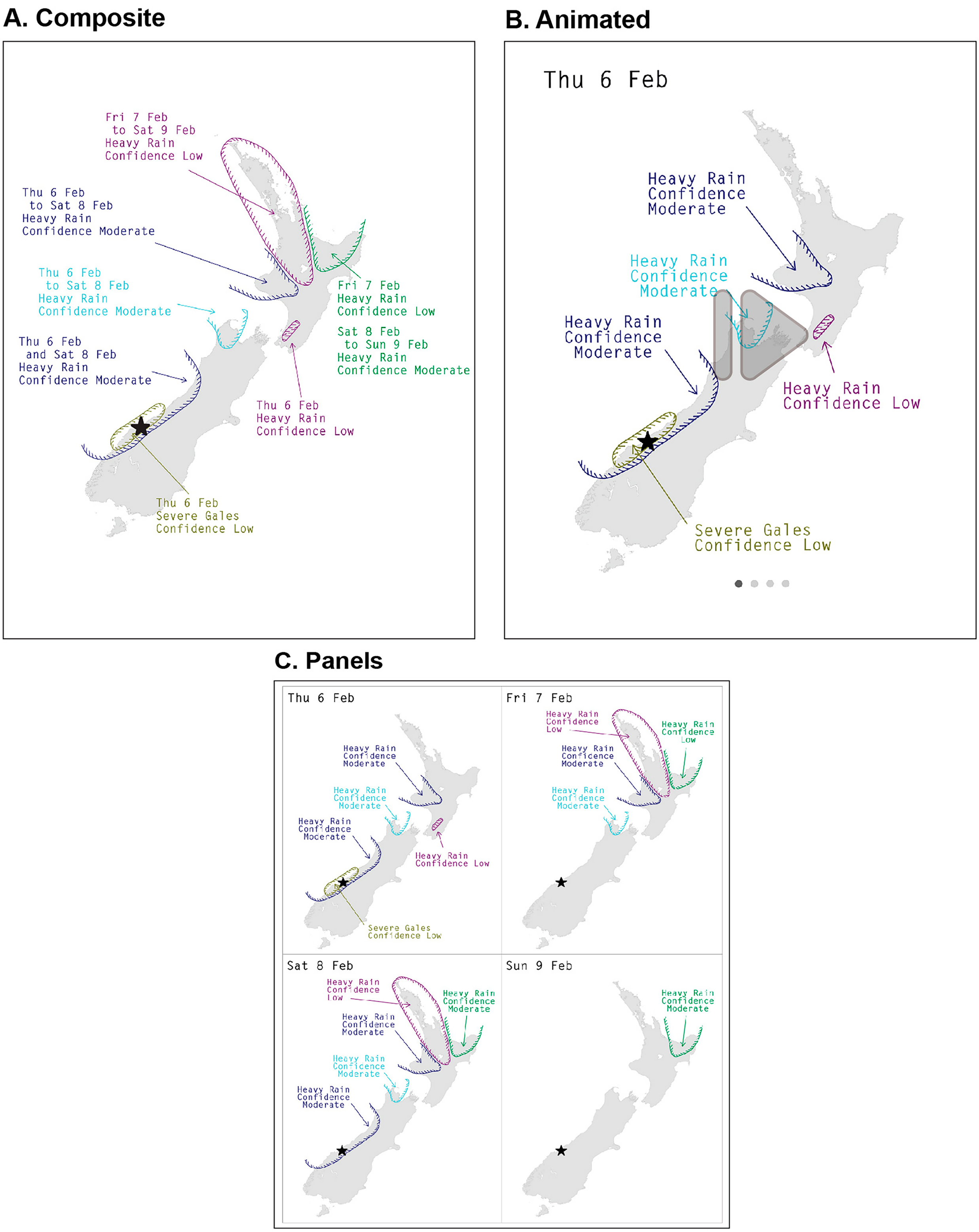

FIG. 2. The (a) composite, (b) animated, and (c) panel maps modified from the MetService Severe Weather Outlook control (Fig. 1) and used in the time frame visualization experiment (experiment A).

one at a time for 3 s each in a loop (similar to the NOAA SPC Convective Outlooks day 4–8 graphic; https://www.spc.noaa.gov/products/exper/day4-8/), and 3) a paneled outlook (univariate) that shows each day presented in an individual panel side by side.

Understanding the severe weather forecast on each day of an extended forecast and its associated uncertainties is important for planning and decision-making across the outlook time frame. For example, it could inform advance planning about whether to carry out an activity on day 3 or 4 of the forecast window. Previous studies have found that perception of likelihood and uncertainty can vary across multiday forecast windows, but it is unclear how visual representation of the multiday time frame influences this process. People anticipate uncertainty in weather forecasts, but their perception of the uncertainty changes for longer lead times (Morss et al. 2008; Lazo et al. 2009; Joslyn and Savelli 2010). For example, Joslyn and Savelli (2010) found that a forecast of rain or snow for the next day was deemed much more likely to occur than a forecast for rain or snow 3 days out. This discrepancy between perceived likelihood and forecast likelihood may arise due to individuals overlaying the forecast with their mental model of how the weather "works," such as its likelihood distribution through time (see, e.g., Doyle et al. 2020), similar to the "base rate effect" (Wallsten et al. 1986; Weber and Hilton 1990; Patt and Dessai 2005; Visschers et al. 2009). Such mental models and biases could play a role in the way audiences interpret outlooks over multiple days in the future and how they perceive and respond to the forecasts. We therefore also explore how perceptions of likelihood, level of concern, and decisions change across the time window of the outlook and whether or not this is affected by the way in which days are shown. We test the following null hypotheses about time visualization approaches, comparing the composite outlook with outlooks with each day of weather represented separately (animated and paneled):

- H1: There is no relationship between the way that days are presented on the outlook (composite or separate) and the accuracy of reporting the days on which severe weather will occur.
- H2: There is no relationship between the way that days are presented on the outlook (composite or separate) and the reported likelihood of severe weather occurring on a day that it is not forecast, when severe weather is forecast to occur on adjacent days.
- H3: There is no relationship between the way that days are presented on the outlook (composite or separate) and the level of concern about driving on a day that severe weather is not forecast, when severe weather is forecast to occur on adjacent days.
- H4: There is no relationship between the way that days are presented on the outlook (composite or separate) and the reported likelihood of changing weekend plans.

c. *Perception and expression of uncertainty*

Uncertainty can also play an important role in how people interpret information from forecast visualizations (Spiegelhalter et al. 2011; Harold et al. 2016). In a systematic review of geospatial visualizations of uncertainty, Kinkeldey et al. (2014) identify five dichotomous categories for uncertainty visualization, including 1) explicit/implicit, 2) intrinsic/extrinsic, 3) visually integral/separable, 4) coincident/adjacent, and 5) static/dynamic. Here, we investigate extrinsic and intrinsic approaches to communicating forecast confidence. Intrinsic approaches alter the existing symbology to represent uncertainty by manipulating visual variables, such as color saturation or hue. Extrinsic approaches add new objects to the display to depict uncertainty, such as symbols or text (see also Bostrom et al. 2008; Kunz et al. 2011; Deitrick and Wentz 2015). The current MetService Severe Weather Outlook graphic uses textual extrinsic expressions of qualitative uncertainty, where "low," "moderate," or "high" confidence is defined as a 20% (or one chance in five), 40% (or two chances in five), or 60% (or three chances in five) likelihood that "the event will actually happen," that is, assigned using a label on the map.

Most previous studies have investigated intrinsic uncertainty depictions, which are considered effective for facilitating an overview of uncertainty (Kinkeldey et al. 2014). Extrinsic expressions, however, have been shown to aid in-depth analysis and support communication of qualitative assessments of uncertainty and have been identified as more suitable for specific locational versus overall uncertainty (Bostrom et al. 2008). Deitrick and Wentz (2015) similarly explored decision-making with "visually integral" and "visually separable" approaches to presenting uncertainty (where the former alters data symbology and the latter overlays uncertainty onto the map with patterns, textures, or shapes). They found that visually integral maps are preferred by decision-makers due to their simplicity, while visually separable images were preferred by researchers and scientists because they could more readily identify and evaluate the uncertainty. This builds on the suggestion by Slocum et al. (2003) that the expressions of uncertainty should consider audience background, as explored by Pang (2008) in a "task-oriented visual mapping" approach that is adapted to different users. The answer to how best to display this visual uncertainty is still unclear due to mixed findings. Thus, Kinkeldey et al. (2014) recommend more empirical studies on how visualization of uncertainty influences decision-making.

Joslyn and Savelli (2010) found that audiences are aware of the inherent uncertainty in meteorological forecasts but are prone to making unjustified assumptions about forecasting biases and uncertainty. For example, participants in their study consistently interpreted an overforecasting bias (perceiving the forecast to be more severe than reality) for extreme weather and an underforecasting bias (perceiving the forecast to be less severe than reality) for rain. Clear communication of uncertainty may be necessary to balance such biases in the interpretation of the forecast (Joslyn and Savelli 2010). Here, we compare comprehension and intended decision-making between extrinsic text expressions of confidence, with the most commonly used intrinsic expression—variations saturation (Kinkeldey et al. 2014; Fig. 3).

We also explore underlying understandings and interpretations of the concept of confidence in relation to this outlook through the following open-ended question: "What do you think the different levels of 'confidence' in the forecast mean?" The World Meteorological Organization (WMO) recognizes "confidence" as

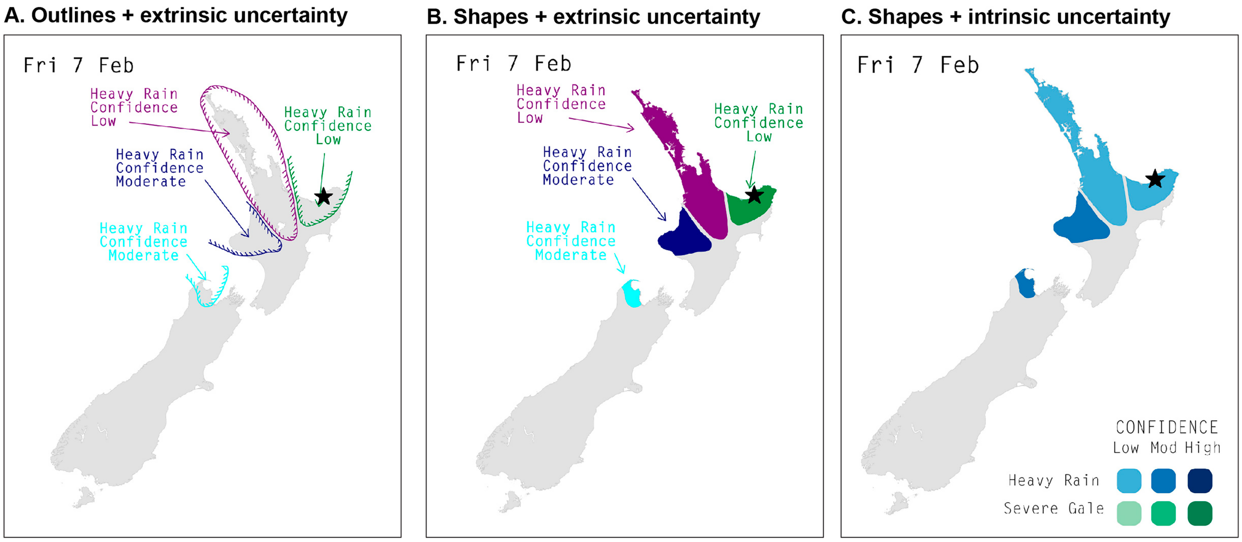

FIG. 3. The (a) outlines + extrinsic, (b) shapes + extrinsic, and (c) shapes + intrinsic modified from the MetService Severe Weather Outlook control (Fig. 1) and used in the perception and expressions of uncertainty experiment (experiment B).

a scale of uncertainty that represents a reliability measure (Gill et al. 2008). The Intergovernmental Panel on Climate Change (IPCC) defines "level of confidence" as the subjective expression of uncertainty based on the degree of understanding in the expert community and degree of agreement across findings and the amount of evidence to support those findings (Mastrandrea et al. 2010). These technical definitions highlight the nuance and complexity of uncertainty terminology and definitions [see review in Doyle et al. (2019)]. While communication of uncertainty is critical for informed decision-making with modeled scientific data, such as severe weather forecasts, the prevalence of many different metrics in the uncertainty lexicon, such as "chance," "probability," "confidence," and "likelihood," can obfuscate what they refer to, especially for nonspecialist audiences [see review in Hudson-Doyle et al. (2018)]. Here, we aim to better understand how people interpret "confidence" in the Severe Weather Outlook context to better understand the role this plays in making inferences from outlook content.

Visual variables in the forecast graphic design, such as shape, hue, size, and texture of forecast areas, can also affect perception of uncertainty by influencing cognitive and emotional responses to the information (e.g., MacEachren 1995; Hegarty 2011; Severtson and Vatovec 2012; Severtson and Myers 2013; Ash et al. 2014; Patterson et al. 2014; Thompson et al. 2015; Padilla et al. 2018). Visual saliency can affect visual attention to mapped data and influence perceived levels of threat (Grant and Spivey 2003; Clive et al. 2021). For example, forecast areas that capture attention through bold shading, bright color, or large size may be associated with higher perceptions of likelihood or degrees of threat. Previous studies have investigated the use of outlines and shaded areas to express probabilistic warning forecasts and found differences in inferences and decision-making between these approaches (e.g., Ash et al. 2014; Thompson et al. 2015; Cheong et al. 2016). Similarly, a recent study by Ernst et al. (2021) found that the colors and language used in NOAA SPC Convective Outlooks in the United States were misinterpreted by some audiences. Here, we compare approaches for representing forecast areas with different visual saliency (outlines vs solid shapes) to explore effects on perception and interpretation of uncertainty (Fig. 3). We test the following four null hypotheses about perceptions and expressions of uncertainty:

- H5: There is no relationship between the forecast area style (outlines or shapes) and the estimated percent confidence in the severe weather occurring.
- H6: There is no relationship between the expression of uncertainty (extrinsic or intrinsic) and the estimated percent confidence in the severe weather occurring.
- H7: There is no relationship between the expression of uncertainty (extrinsic or intrinsic) and the range of estimates of percent confidence in severe weather occurring for days when confidence is "low" and "moderate."
- H8: There is no relationship between the expression of uncertainty (extrinsic or intrinsic) and the day chosen to take a walk.

d. Icons

Graphical representations, such as icons and symbols, are increasingly prevalent in modern mobile applications for weather forecasts. Symbols such as a sun, a cloud, or raindrops are commonly used as icons in forecast products around the globe to quickly communicate simple weather information (Zabini 2016). Studies show that combining icons with text information is preferred among users (Carr et al. 2016). However, despite their prevalence, icons are often ineffective at clearly illustrating specific weather phenomena (Zabini et al. 2015; Zabini 2016; Reed et al. 2022). People also tend to assign varying amounts of uncertainty to deterministic weather icons (Sivle et al. 2014; Zabini et al. 2015). While the connotations and nuanced meaning of

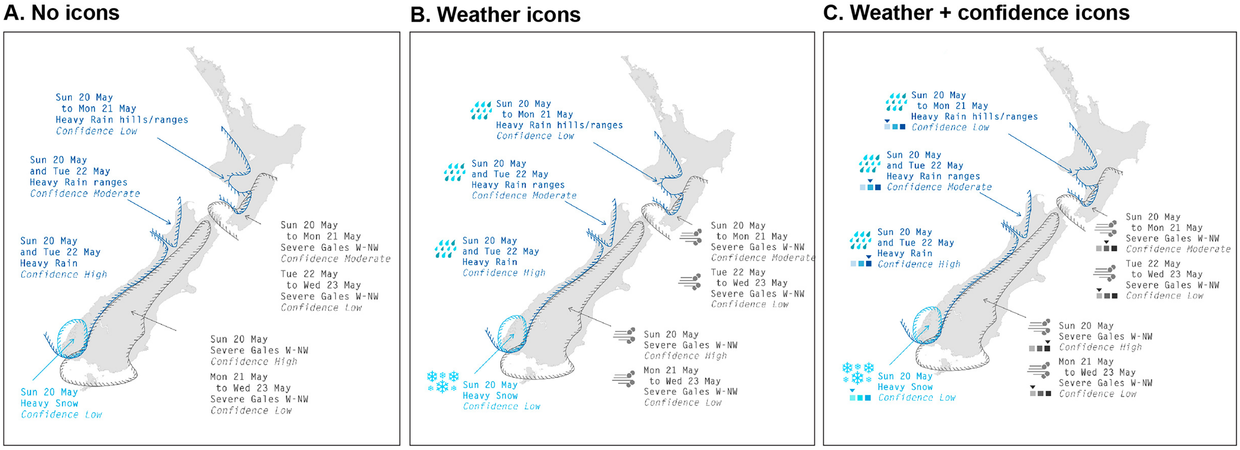

FIG. 4. The (a) icons, (b) weather icons, and (c) weather + confidence icons maps modified from the MetService Severe Weather Outlook control (Fig. 1) and used in the icon experiment (experiment C).

weather icons is complex and experiential for many forecast users (Sivle et al. 2014; Reed and Senkbeil 2020), icons may play a role beyond phenomenon identification, especially over long-term outlook time frames. Weather icons may serve as visual cues for "chunking" the weather processes by phenomenon to understand overall distributions and frequency of severe weather across the region without having to read text descriptions for each forecast area. To explore the potential value of weather icons in recognizing and distinguishing large-scale weather patterns and trends, we added gale, heavy rain, and snow icons used in MetService's daily weather forecast to the Severe Weather Outlook for two of the three experiment groups (Fig. 4).

Icons have also been shown to assist in the interpretation of uncertainty and can lead to increased decision performance relative to numerical or textual representations alone (Bisantz et al. 2005, 2011). However, much of the existing work on uncertainty icons focuses on intrinsic methods for representing uncertainty, through adjusting visual variables of the icon display, such as saturation, resolution, or shape of the icon (e.g., MacEachren 1992; Bisantz et al. 2005, 2011; Kinkeldey et al. 2014; Kolbeinsson et al. 2015). Here, we explore the potential value of icons to communicate uncertainty extrinsically. For one experiment group, we therefore added both weather icons and an icon to represent the confidence, or uncertainty, where an arrow points to one of three increasingly saturated bars representing low, medium, and high confidence (Fig. 4). We test the following null hypotheses about the use of icons:

- H9: There is no relationship between the icons shown on the outlook and the accuracy in reporting the number of different weather phenomena occurring during the outlook period.
- H10: There is no relationship between the presence of confidence icons and the accuracy in responding to statements about forecast likelihood.
- H11: There is no relationship between the presence of confidence icons and the accuracy in comparing the likelihood of experiencing severe weather between the North and South Islands.
- H12: There is no relationship between the presence of icons and the reported ease of reading information from the outlook.

2. Methodology

a. Study design and sampling

We employed a purposive sampling online survey approach in order to reach a large national sample of active weather forecast users. Although purposive sampling does not enable us to test a representative sample of the population of NZ, this non-probability convenience sampling approach allowed us to target a subset of the primary users of the MetService Severe Weather Outlook in order to increase the relevancy and reliability of findings for this audience. The survey was advertised on the MetService web site and its social media channels (e.g., Twitter, Facebook) and was shared with relevant special interest groups in order to reach audiences who actively use and seek long-term weather forecasts. Data were collected using the Qualtrics online survey platform, and the survey was open for participation for 2 months from mid-October to mid-December 2019.

A between-group mixed-methods study design was adopted, in which participants were randomly assigned to one of three groups in which they answered the same set of multiple-choice and open-ended questions using different maps. Each group saw different modified versions of the MetService Severe Weather Outlook, and responses were compared across groups. The survey was designed around the following three successive experiments in which variables were changed across groups: 1) time frame visualization, 2) perception and expression of uncertainty, and 3) icon inclusion (Table 1). Questions explored comprehension (reporting the correct answer), perceptions of likelihood (Likert scales and percentages), and self-reported ease of reading (Table A1). In the rest of this section, the name of each group within experiments A–C is italicized on first use for clarity.

The outlook design used in experiments A and B was based on a MetService Severe Weather Outlook published on 9 May 2018 that included 4 days (11–14 May) of forecasts for

TABLE 1. Experimental groups and outlook variables.

Expt	Group 1	Group 2	Group 3
A. Time frame visualization	Composite[a]	Animated	Panel
B. Expressions of uncertainty	Outlines + extrinsic	Shapes + extrinsic	Shapes + intrinsic
C. Icons	No icons	Weather icons	Weather + confidence icons

[a] MetService map (control).

heavy rain and severe gales across the country with low and moderate confidence (Fig. 1). The dates were changed on the outlook to reflect a summer public holiday weekend in New Zealand that is popular for traveling and planning outdoor activities (Waitangi Day long weekend). For the experiments, participants were asked to imagine that they were planning activities several days out from the Waitangi Day long weekend (6–9 February) and then respond to a series of questions about their planning while viewing the modified outlooks. In experiment A, the time frame visualization was modified such that the *composite* group saw one composite map graphic representing all 4 days of weather (control map; Fig. 2a); the *animated* group saw one animated map graphic [graphics interchange format (GIF)], which cycled in a loop showing each day (1–4 successively) for 3 s at a time and could not be paused by the user (Fig. 2b); and the *panel* group saw a four-paneled map graphic that juxtaposed the 4 days beside each other (Fig. 2c). Participants used the maps to respond to questions about the weather and travel plans for a fictional campsite at a starred location that was forecast to experience heavy rain and gales on Thursday and heavy rain on Saturday but no severe weather on Friday and Sunday.

The paneled visualization from experiment A (panel group) was adopted for experiment B. In experiment B, two groups saw outlooks in which the type of weather and confidence level were extrinsically written using text. For the *outlines + extrinsic* group, forecast areas were represented by hashed outlines of random color with a text label defining the confidence level (current MetService approach; Fig. 3a). For the *shapes + extrinsic* group, forecast areas were represented by solid shapes of random color with a text label defining the confidence level (Fig. 3b). The *shapes + intrinsic* group was shown an outlook in which the forecast areas were represented by solid shapes, but a shape's hue represented the type of weather, and the saturation represented the confidence level intrinsically (Fig. 3c). Participants used the maps to respond to questions about confidence and planned activities for a campsite at a starred location that was forecast to experience heavy rain on Friday (low confidence), Saturday (moderate confidence), and Sunday (moderate confidence).

The composite-style outlook design used in experiment C was based on a MetService Severe Weather Outlook published on 18 May 2018 that included 4 days (20–23 May) of forecasts for heavy rain, severe gales, and heavy snow across the country with low, moderate, and high confidence. The forecast area hue represented the type of weather (blue, heavy rain; light blue, heavy snow; gray, severe gales). For all groups, the type of weather was described in a written text label pointing to the forecast area (Figs. 4a–c). The *no icons* group had no icons shown on the map. The *weather icons* group saw an outlook that included weather icons (Fig. 4b). The *weather and confidence icons* group saw an outlook that included both weather icons and icons that extrinsically represented the level of confidence (Fig. 4c). Participants then responded to questions about the types of weather occurring across the outlook period, forecast confidence, and likelihood.

At the end of the survey, all participants responded to a set of multiple-choice questions about demographics; familiarity, confidence, and use of weather forecasts and the MetService Severe Weather Outlook; and an open-ended question about the perceived meaning of the "confidence" levels.

b. Statistical analysis

We used IBM SPSS Statistics version 24.0.0.1 to carry out nonparametric statistical analyses on our null hypotheses and report findings at the $p < 0.05$ significance level. We use the χ^2 test for equality of proportions to determine if participant characteristics are equally distributed between experiment groups (e.g., to identify if one group reported being more familiar with the outlooks than the other groups). We use the Pearson's χ^2 test to test for associations in categorical variables across experiment groups and report φ_c as a test of the strength of the association. We use the Mann–Whitney U test to compare differences between independent continuous variables (i.e., percent confidence reported) between groups, and we use the Kruskal–Wallis test to test for differences in the continuous variables (i.e., changes in percent confidence reported for "low" and "moderate") across all three groups. Where noted, some groups were combined for analysis. For example, the paneled and animated approaches of time representation were grouped together as "separate" days for comparison with the composite days group in some instances to assess differences between bivariate and univariate approaches. We use the Shapiro–Wilks test to test for normality before applying nonparametric tests. A content analysis (Creswell 2013) was carried out on text responses to open-ended questions to identify the presence and frequency of words, themes, and concepts raised by participants. Inductive codes were recognized from the data, through a process of reading and interpreting the data to identify key concepts and to further explore participants' perspectives and understanding of the design options.

3. Results

a. Participant overview

The online survey was open for 2 months, and a total of 417 people participated, with 369 completing the survey in

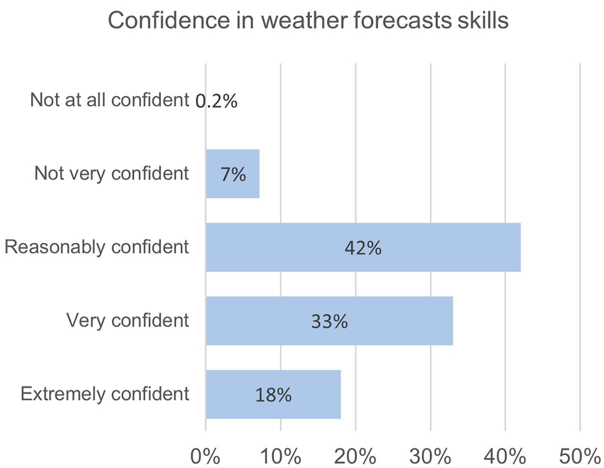

FIG. 5. Self-reported confidence in skills for using and reading weather forecasts.

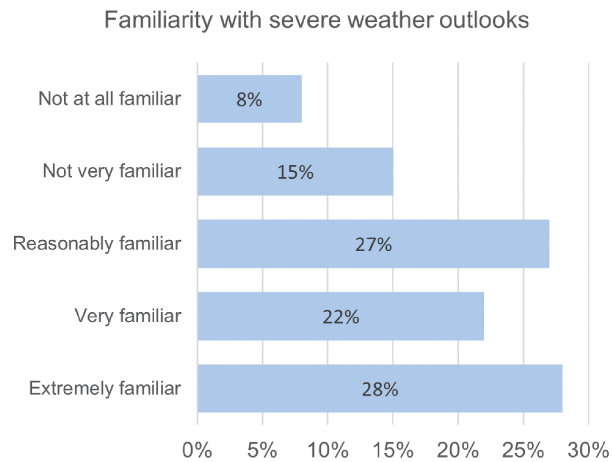

FIG. 6. Self-reported familiarity with Severe Weather Outlook graphics.

full. Of those who responded to a question about language ($n = 396$), 100% reported being fluent in English. Of those who responded to questions about age and education ($n = 369$), a majority were between the ages of 20 and 69 [93% ($n = 343$)] and had completed a university degree [68% ($n = 269$)]. When participants were asked about where they spent most of their time, responses ($n = 417$) were distributed across all 15 regions of NZ, with higher representation from regions with large urban centers, such as the capital city of Wellington [42% ($n = 177$)], Auckland [13% ($n = 56$)], and Canterbury [12% ($n = 48$)]. Of the 417 respondents, most were long-term residents, having resided in NZ for 20 years or more [77% ($n = 321$)] or 10–19 years [11% ($n = 44$)].

More than one-half of the 417 respondents reported high confidence in their skills for "reading and using weather forecasts" [51% ($n = 211$)], with 18% ($n = 75$) "extremely confident" and 33% ($n = 136$) "very confident" (Fig. 5). A large portion reported moderate confidence, with 42% ($n = 175$) "reasonably confident," and 7.4% ($n = 31$) reported low confidence, with 7.2% ($n = 30$) "not very confident" and 0.2% ($n = 1$) "not at all confident." A majority [88% of the 417 respondents ($n = 367$)] reported using weather forecasts regularly (daily to weekly), 9.4% ($n = 39$) occasionally (e.g., planning trips and occasions), and 2.4% ($n = 10$) rarely (e.g., in storm events); 0.2% ($n = 1$) reported that they never used weather forecasts.

One-half of the 417 respondents also reported high levels of familiarity with the MetService Severe Weather Outlook [50% ($n = 206$)], with 28% ($n = 116$) "extremely familiar" and 22% ($n = 90$) "very familiar" (Fig. 6). Less than one-third reported moderate levels of familiarity, with 27% ($n = 111$) "reasonably familiar," and low levels of familiarity [24% ($n = 100$)], with 15% ($n = 62$) "not very familiar," 8.6% ($n = 36$) "not at all familiar," and 0.5% ($n = 2$) "not sure." There were no observed differences in the level of confidence with weather forecasts [$\chi^2(4) = 7.66$; $p = 0.105$; $\varphi_c = 0.096$] or familiarity with the outlooks [$\chi^2(4) = 5.39$; $p = 0.249$; $\varphi_c = 0.080$] between experiment groups at the $p = 0.05$ level. At the start of the survey, 61% of 417 respondents reported a high likelihood of using the MetService Severe Weather Outlook for planning activities such as family outings, with 32% ($n = 134$) reporting they were "extremely" and 29% ($n = 121$) "very" likely to use it. Similarly, most of the 417 respondents [69% ($n = 286$)] reported that, overall, the Severe Weather Outlook was "extremely" or "very useful" as a tool for planning ahead, with 28% ($n = 115$) responding it was "reasonably useful" and 4% ($n = 16$) responding it was "not very" or "not at all useful."

b. Experiment A: Time frame visualization

Four null hypotheses were tested in experiment A to explore the effect of time frame visualization on comprehension, perceived likelihoods, and behavioral intent, using an outlook that forecast heavy rain for Thursday and Saturday, but not Friday.

- H1: There is no relationship between the way that days are presented (composite or separate) and the accuracy of reporting the days on which severe weather will occur.

Experiment A failed to reject H1. When participants were asked to report which days were forecast to experience severe weather, there were no observed differences in the proportion of correct responses between the composite group and the groups that saw days presented separately (animated and panel groups) [$\chi^2(2) = 1.72$; $p = 0.423$; $\varphi_c = 0.064$]. Of the 417 respondents, a majority of participants in each group were able to correctly identify that the campsite was forecast to experience severe weather only on Thursday and Saturday (composite: 62%, with $n = 86$ of 139 respondents; animated: 63%, with $n = 88$ of 140 respondents; and panel: 69%, with $n = 43$ of 138 respondents).

- H2: There is no relationship between the way that days are presented (composite or separate) and the reported likelihood of severe weather occurring on a day that it is not forecast, when severe weather is forecast to occur on adjacent days.

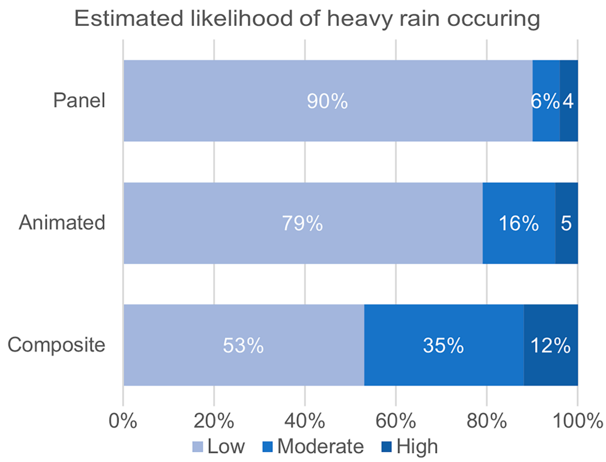

FIG. 7. Estimated qualitative likelihood of heavy rain occurring on Friday.

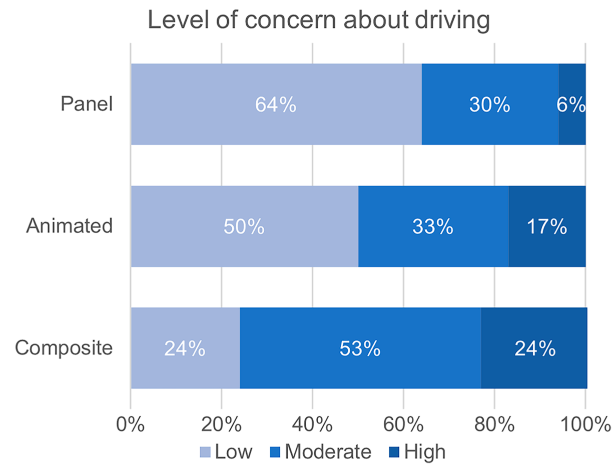

FIG. 8. Reported level of concern about driving on roads on Friday.

H2 was rejected. Across all respondents ($n = 417$), participants in the composite group ($n = 139$) were more likely to respond with a high or moderate likelihood that heavy rain would occur on Friday 7 February at the campsite than participants in the groups that saw weather for each day presented separately (animated and panel) [$\chi^2(4) = 52.3$; $p < 0.01$; $\varphi_c = 0.250$], with 12% ($n = 17$) reporting that heavy rain was "extremely" or "very" likely to occur and 35% ($n = 48$) responding that it was "reasonably" likely to occur (Fig. 7). The highest portion of correct responses (low likelihood) occurred among members of the panel group ($n = 138$), of whom 90% ($n = 124$) reported that heavy rain was "not very" or "not at all" likely to occur.

- H3: There is no relationship between the way that days are presented (composite or separate) and the level of concern about driving on a day that severe weather is not forecast, when severe weather is forecast to occur on adjacent days.

H3 was rejected. Across all respondents ($n = 416$), participants in the composite group ($n = 139$) were more likely to respond with a high or moderate level of concern about driving on the roads near the campsite on Friday 7 February [$\chi^2(4) = 50.06$; $p < 0.01$; $\varphi_c = 0.245$], with 24% ($n = 33$) reporting that they would be "extremely" or "very" concerned, 53% ($n = 73$) reporting that they would be "reasonably" concerned, and 24% ($n = 33$) responding "not very" or "not at all" concerned (Fig. 8). Low levels of concern ("not very" or "not at all") were more frequently reported among most participants in the groups that saw weather for each day presented separately (animated: 50%, with $n = 69$ of 139 respondents, and panel: 64%, with $n = 88$ of 138 respondents).

- H4: There is no relationship between the way that days are presented (composite or separate) and the reported likelihood of changing weekend plans.

H4 was rejected. Across 417 respondents, participants in the composite ($n = 139$) and animated ($n = 140$) groups were more likely to respond with a high or moderate likelihood of changing plans based on the forecast relative to the panel group ($n = 138$) [$\chi^2(4) = 10.269$; $p = 0.036$; $\varphi_c = 0.111$]. The panel group was more likely to report a low likelihood of changing plans [22% ($n = 30$)] relative to the composite group [11% ($n = 15$)] or animated group [14% ($n = 19$)] (Fig. 9). The panel-group participants were also less likely to report a high likelihood of changing plans [40% ($n = 55$)] relative to the composite [56% ($n = 78$)] and animated [51% ($n = 71$)] groups. When optionally describing how they would change plans, respondents across all groups generally described that they would either cancel their plans completely or delay the start of their trip to Friday to miss inclement weather on Thursday. Overall, the composite outlook was associated with higher perceived likelihoods of severe weather, levels of concern, and intent to change plans than the outlooks that presented information grouped into separate days (animated and panel style outlooks).

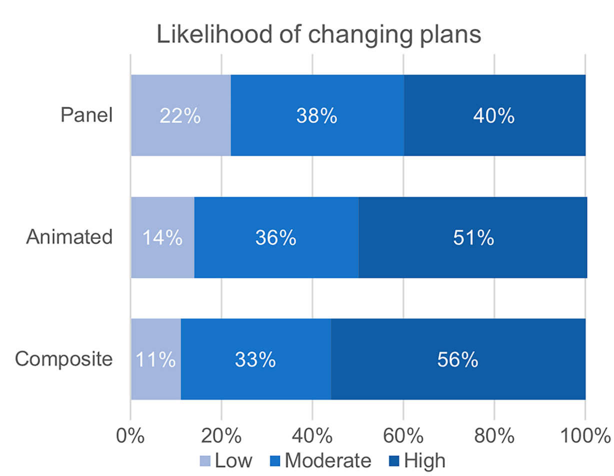

FIG. 9. Reported qualitative likelihood of changing plans based on the outlook forecast.

TABLE 2. Mean, median, and mode for estimated confidence across groups.

Group	Mean	Median	Mode	Std dev
Confidence (0%–100%) in heavy rain on Friday 7 Feb ("low")				
Outlines + extrinsic	30	25	20	19
Shapes + extrinsic	35	29	20	24
Shapes + intrinsic	40	31	30	24
Confidence (0%–100%) in heavy rain on Saturday 8 Feb ("moderate")				
Outlines + extrinsic	58	58	50	17
Shapes + extrinsic	59	60	50	19
Shapes + intrinsic	69	69	100	21

Panel-group participants reported the highest ease of reading information from the outlook [$\chi^2(4) = 10.276$; $p = 0.036$; $\varphi_c = 0.111$]. The panel-group respondents ($n = 138$) were more likely to report a high level of ease ["extremely" or "very" easy; 62% ($n = 86$)] compared to the composite group [50% ($n = 69$ of 139 respondents)] or animated group [45% ($n = 63$ of 140 respondents)]. Low levels of ease ("not very" or "not at all" easy) were most frequently reported among the animated group [16% ($n = 23$ of 140 respondents)].

c. *Experiment B: Perception and expression of uncertainty*

Four null hypotheses were tested in experiment B to explore how representation of the forecast area and expressions of uncertainty affect perceptions and estimates of forecast confidence.

- H5: There is no relationship between the forecast area style (outlines or shading) and the estimated percent confidence in the severe weather occurring.

H5 was rejected. The mean estimated percent confidence among those who saw areas displayed as solid shapes ($n = 264$) was higher than that among those who saw outlines ($n = 133$; independent-samples Mann–Whitney U test) for both Friday (low confidence; $U = 20\,486$; $p = 0.001$) and Saturday (moderate confidence; $U = 21\,108$; $p = 0.001$; Table 2; Fig. 10).

- H6: There is no relationship between the expression of uncertainty (extrinsic or intrinsic) and the estimated percent confidence in the severe weather occurring.

H6 was rejected. Extrinsic uncertainty (e.g., "low"/"moderate"; $n = 264$) was associated with lower mean estimates of confidence than intrinsic uncertainty (hue saturation; $n = 133$; independent-samples Mann–Whitney U test) for both Friday (Q12; low

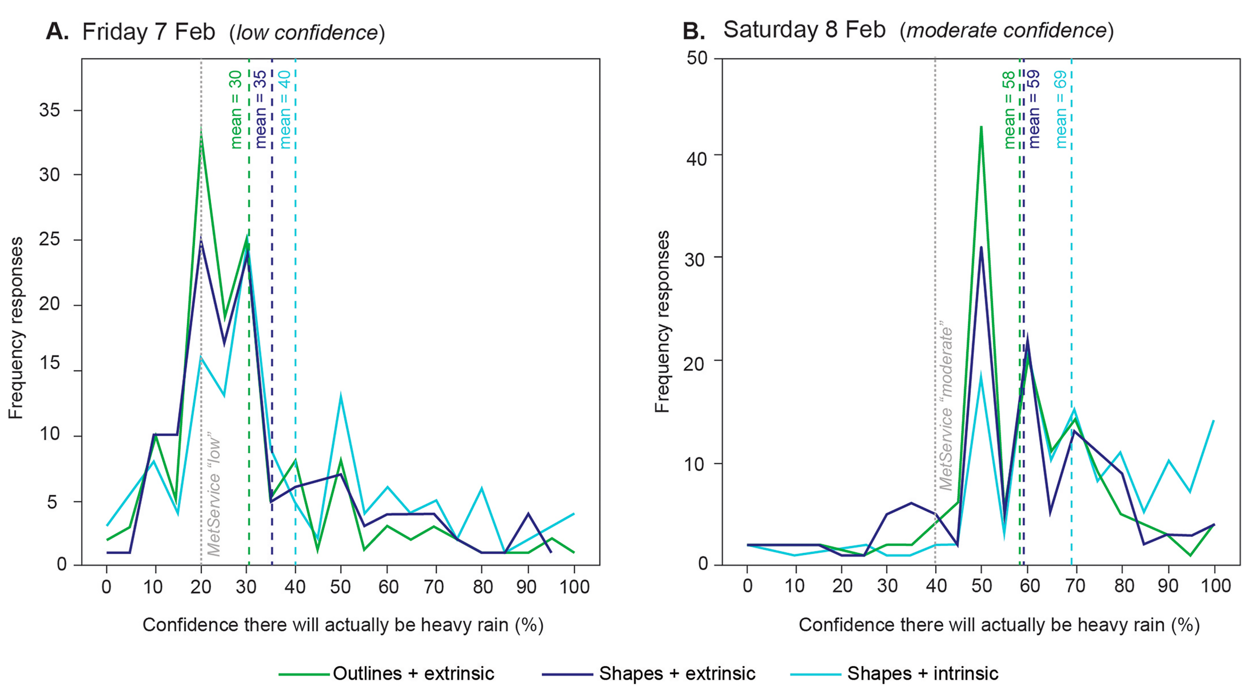

FIG. 10. Reported percent confidence that heavy rain will actually occur at the campsite on Friday (when heavy rain has low forecast confidence) and Saturday (when heavy rain has moderate forecast confidence).

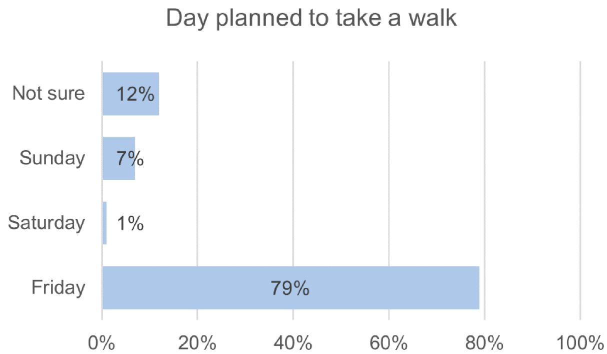

FIG. 11. Reported day chosen to take a walk as based on the outlook forecast.

confidence; $U = 21\,322$; $p < 0.01$) and Saturday (Q13; $U = 23\,354$; $p < 0.01$).

- H7: There is no relationship between the expression of uncertainty (extrinsic or intrinsic) and the range of estimates of percent confidence in severe weather occurring for days when confidence is "low" and "moderate."

Experiment B failed to reject H7. There were no significant differences in the range between estimates for Friday ("low" confidence) and Saturday ("moderate" confidence) between groups that saw uncertainty displayed extrinsically and intrinsically [independent-samples Kruskal–Wallis test; $\chi^2(2) = 0.340$; $p = 0.843$].

- H8: There is no relationship between the expression of uncertainty (extrinsic or intrinsic) and the day chosen to take a walk.

Experiment B failed to reject H8. Across all responses ($n = 397$), a majority of participants [79% ($n = 315$)] responded they would choose to take a walk on Friday, the day with the lowest confidence in the severe weather forecast for the area, and there was no significant difference in the day chosen and the outlook style viewed [$\chi^2(4) = 4.89$; $p = 0.299$; $\varphi_c = 0.79$; Fig. 11].

Participants in the shapes + extrinsic group ($n = 132$; Fig. 3b) reported the highest ease of reading information from the maps [$\chi^2(4) = 9.54$; $p = 0.049$; $\varphi_c = 0.109$]. Participants in the shapes + extrinsic group were more likely to report a high level of ease ["extremely" or "very" easy; 57% ($n = 75$)] relative to participants in the outlines + intrinsic group [$n = 133$; 47% ($n = 63$)] or participants in the shapes + intrinsic group [$n = 133$; 50% ($n = 67$)] group. Low levels of ease ("not very" or "not at all" easy) were most frequently reported among the shapes + intrinsic group [17% ($n = 23$); Fig. 3c].

d. *Experiment C: Icons*

Four null hypotheses were tested in experiment C to explore how the presence of icons affects performance and ease of reading the outlooks.

- H9: There is no relationship between the icons shown on the outlook and the accuracy in reporting the number of different weather phenomena occurring during the outlook period.

H9 was rejected. The addition of weather icons increased the frequency of correct responses among participants who responded to a question asking them to report how many weather phenomena were forecast to occur over the outlook period [$n = 378$; $\chi^2(1) = 6.37$; $p = 0.012$; $\varphi = 0.156$]. While 78% ($n = 96$) of respondents in the no icon group ($n = 124$) correctly identified that three different weather phenomena (gales, rain, snow) were forecast to occur, 90% ($n = 115$) of participants in the weather icons group ($n = 128$) reported correct answers. However, respondents in the weather + confidence icons group did not show a significant improvement over the no icons group, with 82% ($n = 103$) of the weather + confidence icons group ($n = 126$) reporting the correct answer of three [$\chi^2(1) = 0.484$; $p = 0.487$; $\varphi = 0.044$].

- H10: There is no relationship between the presence of confidence icons and the accuracy in responding to statements about forecast likelihood.

Experiment C failed to reject H10. When presented with several statements about forecast intensity, accuracy, and likelihood, there were no significant differences in participants' ability to identify the "true" statement between groups [$\chi^2(2) = 0.005$; $p = 0.998$; $\varphi_c = 0.004$]. The correct answer, "the likelihood is forecast to vary between days," was chosen by less than 40% of participants ($n = 151$) across all groups ($n = 378$) as follows: no icons [36% ($n = 44$ of 124 respondents)], weather icons [37% ($n = 47$ of 128 respondents)], and weather + confidence icons [37% ($n = 46$ of 126 respondents)].

- H11: There is no relationship between the presence of confidence icons and the accuracy in comparing the likelihood of experiencing severe weather between the North and South Islands.

Experiment C failed to reject H11. Across participants in all groups ($n = 378$), there were no significant differences in the ability to correctly identify that the South Island was more likely to experience severe weather than the North Island [$\chi^2(2) = 0.974$; $p = 0.615$; $\varphi_c = 0.051$]. A majority of participants reported the correct response across all groups as follows: no icons [79% ($n = 98$ of 124 respondents)], weather icons [74% ($n = 94$ of 128 respondents)], and weather + confidence icons [79% ($n = 99$ of 126 respondents)].

- H12: There is no relationship between the presence of icons and the reported ease of reading information from the outlook.

H12 was rejected. Participants in the no icons group reported the highest ease of reading information from the outlook [$\chi^2(4) = 25.5$; $p < 0.01$; $\varphi_c = 0.184$; Fig. 12].

e. *Interpretations of "confidence"*

When asked, "What do you think the different levels of confidence in the forecast mean?" participants ($n = 329$)

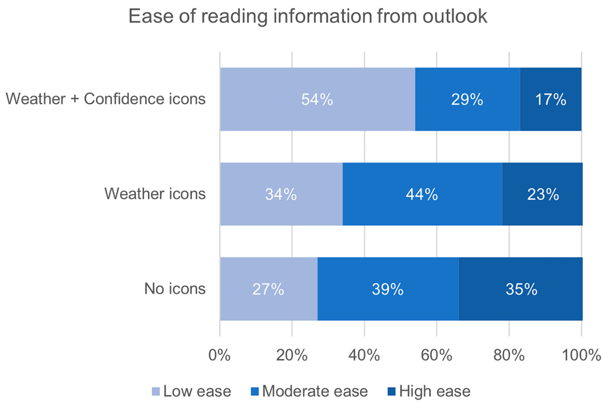

FIG. 12. Self-reported qualitative level of ease of reading information from the outlook.

demonstrated a wide range of latent interpretations of the word "confidence," with the most dominant themes relating to likelihood, statistics, probability, or chance, followed by a consideration of the accuracy of the data and the agreement between models, forecasts, or forecasters and then that the participants were confused, did not know, or believed confidence reflects a subjective best guess or hedging (Table 3). The most common themes related to participants translating "confidence" into either likelihood ($n = 156$) or occurrence ($n = 133$) or both ($n = 83$), translating it directly to "the likelihood of occurrence," for example, "the likelihood that the severe weather will occur" or "confidence in that weather occurring at that time." This indicates that nearly a quarter of participants were interpreting the concept the way intended by MetService. Relating the confidence to statistics and probability, or possibility and chance, was raised by less than a fifth of participants ($n = 61$ and 37, respectively), for example, a "statistical chance of happening" or "possibility of the forecast being correct" or "a longwinded way of saying 'chance.'" A small number of participants ($n = 10$) referred to the confidence being subjective and a best guess, based on gut feelings or opinion, for example, "a 'best guess' at the likelihood of a particular weather pattern."

In the next dominant themes, participants considered the forecasts, predictions, data, models, and calculations ($n = 69$) and the confidence in those ($n = 33$), for example, "confidence the forecasters have in their prediction being true" and "how precise the info is." A number of participants inferred that confidence either was related to the uncertainty and disagreement among advice and forecasters ($n = 45$) or reflected the truth and reliability of the forecast or accuracy of data ($n = 41$), for example, "how likely the forecast is to match actual weather" and "confidence the forecasters have in their prediction being true" and "a measure of the uncertainty (divergence) in the models." These interpretations relate closely to how confidence is defined by agencies such as the WMO (Gill et al. 2008) and the IPCC (Mastrandrea et al. 2010).

Thirty-five respondents also indicated they were confused and did not know what "confidence" meant or did not like its use, for example, "a clumsy way of expressing percentage of something being likely" and "I always struggle with the confidence levels; it's very hard to assess how much to trust it." This latter quote also highlights an additional theme raised, that is, trust, either the trust the forecasters have in the data and models or participants' own trust in the forecasts presented to them.

TABLE 3. Summary of thematic analysis of participant descriptions of confidence.

Theme (master code)	Description (discussion or mentions)	Frequency
Occurrence	Mentions "occurrence"	156
Likelihood	Mentions "likelihood"	133
Forecasts, predictions	Forecasts, predictions by models or forecasters	69
Data, models, calculation	Aspects of models, data, data quality, estimates, or calculations	33
Statistics, probability	Statistics, percentages, or probabilities	61
Possibility, chance	Chance or possibility	37
Un-/certainty, dis-/agreement	Agreement or disagreement of advice of forecasters, model agreement	45
Truth, accuracy	Truth, reliability of forecast to match reality; discusses accuracy of forecast or data	41
Confusing, do not know	Confusing, does not like them, finds them difficult, clumsy	35
Level of belief, sureness	Depends upon belief of forecasters; their level of "sureness"	15
Subjective, guess	Subjectivity, guessing, or gut feelings	10
Weather phenomena	Specifically discusses weather phenomena (e.g., rain) or that it depends on weather type	23
Risk, impacts, intensity	Relates their understanding to risk, impacts, severity, and intensity	14
Path, space, time	Change in path direction, location, etc., or discusses times, time frames, days of event, etc.	27
Framing, relevance	Refers to it being a generic or lay framing, or consumer relevance, communicate understanding	7
Warnings, precaution	Discusses warnings or thresholds for warnings, or a precautionary approach	15
Nonoccurrence	Nonoccurrence	11
Variability	Forecast or weather variability	10
Translation	Provides a translation of confidence in some form (to probabilities or likelihood, etc.)	25
Forecaster experience	Depends upon forecasters experience, or forecasters' trust in the forecast and models	3
Participant trust	Participants discuss their trust	3
Map characteristics	Elements of the map, such as color, lines, etc.	1
Not likelihood	Explicitly stating it is not equivalent to likelihood	1

4. Discussion

Severe weather outlooks communicate important information about approaching weather hazards, including their distribution, duration, and associated uncertainty. Visualizing this information in a forecast product requires challenging design decisions, each with competing communication advantages. In this study, we explore how such visualization choices affect the inferences and intended decisions made using the outlooks. We find that variables for time frame visualization, expressions of uncertainty, and representation of the forecast area and phenomena can influence levels of concern, planning decisions, and outlook comprehension. The findings have several implications for supporting clear communication and informed decision-making.

a. Grouping content by day can reduce perceived biases

Outlooks that divide multiday forecast information into a format that presents each day separately may result in more deliberate decision-making that is less influenced by perceived biases. Our results indicate that people who view composite-style maps, which show information for all four days of weather on one map, may incorrectly infer that forecast areas extend into neighboring days of the outlook. In experiment A of our study, this emerged as a perceived underforecasting bias. Participants in the composite outlook group reported higher likelihoods of severe weather occurring on a day when there was no severe weather forecast to occur, despite accurately reporting that there was no severe weather forecast on that day (Fig. 7). They tended to extend the severe weather forecast for Thursday and Saturday into the forecast for Friday (suggesting an assumption that it was underforecast). This perceived bias resulted in higher reported levels of concern about travel on Friday (Fig. 8) and greater likelihood of changing plans for weekend travel (Fig. 9) relative to the groups that saw the information separated into days. This aligns with previous findings that audiences may make unwarranted assumptions about forecasts and uncertainty (Joslyn and Savelli 2010) and provides empirical evidence that forecast visualization choices could play a role in mitigating these types of perceived biases.

While a composite map offers advantages by distilling information about multiple days into a single image, there are tradeoffs associated with this approach. When the information was separated out into individual days, people were less likely to perceive that the outlook was underforecasting the weather for days adjacent to those with severe weather. The visual cue of seeing severe weather forecasts illustrated on some days and not on others may have affected perceptions of forecast accuracy and reliability in the animated and panel groups. Displaying forecast information for each day individually in a univariate approach may help reduce the perceived underforecasting bias associated with composite styles. We found that panels, in particular, have a higher likelihood of being interpreted correctly and were also reportedly easier to read and understand than the animated style. A well-designed interactive map, where the user can choose to view individual days, could be an effective alternative to the panel and animated maps tested in this study. Although they are less transferable across media platforms, interactive maps offer many advantages for engaging map users and empowering them to create a representation that meets their needs (Roth et al. 2017), and future work should investigate this approach.

Across all participant groups, 10%–47% of participants reported moderate to high likelihoods of heavy rain occurring on a day that did not have any heavy rain forecast but was between 2 days for which heavy rain was forecast. This indicates complex perceptions of the timing and evolution of severe weather events, which should be further explored in future work.

b. Visual saliency can affect perceived forecast confidence

Forecast areas represented as solid shapes, which are more vivid and eye catching, or more visually salient, than the outlined forecast areas, were associated with higher estimates of forecast confidence. Participants who viewed forecast areas as shapes estimated higher levels of forecast confidence than participants who viewed forecast areas in outlines of the same color. The solid shapes were associated with higher median and mean percent confidence that the heavy rain forecast would "actually" occur for both "low confidence" and "moderate confidence" forecasts (Fig. 10; Table 2). Similarly, confidence visualized intrinsically, using variations in saturation for the visually salient shapes, resulted in higher median and mean percent confidence estimates than the extrinsic (textual) approaches. Estimates of confidence among participants who saw forecast areas as outlines and/or extrinsic expressions of uncertainty more closely matched the forecasters' intended confidence values. For example, the mean and median values reported for "low" and "moderate" among the outlines and extrinsic groups were closer to the 20% and 40% values to which these words were intended to refer.

Visual saliency is a key variable for capturing attention and driving cognition (Itti and Koch 2000; Cavanagh 2011), and previous studies have shown that it can influence decision-making with maps and graphics (Jarvenpaa 1990; Fabrikant et al. 2010; Clive et al. 2021). Our findings suggest that visual saliency also plays a latent role in how people cognitively assess the uncertainty or confidence in outlook forecasts. The tangible nature of seeing data visualized on a map can result in map readers ascribing higher certainty to the information than warranted (MacEachren 1992; Severtson and Vatovec 2012). Using bold visually salient colors with sharp boundaries to display forecast areas could reinforce this bias by increasing perceived forecast confidence. Shapes also present a number of visualization challenges in the instances of overlapping forecast areas, which may occur in complex forecasts or longer-term outlooks. For example, the area starred in Fig. 2 has overlapping gales and heavy rain on Thursday 6 February. This requires interpretation of whether or not both events are occurring in the overlapping shaded areas and introduces space for divergent interpretations.

In this study, we do not explore the role of color choice in how shape and outlined forecast areas are interpreted. However, we acknowledge that color is widely recognized as having an effect on the way people comprehend and interpret maps (e.g., Robinson 1967; Bryant et al. 2014). We suggest that future iterations of the Severe Weather Outlook develop a consistent and transparent approach to color choice, as people may infer unintended meaning from a random color choice, which could also result in divergent interpretations of the same forecast.

c. *Confidence is informative but confusing*

Our results indicate that participants anticipate uncertainty with outlook forecasts and can effectively use forecast confidence to inform their decision-making, despite an inconsistent interpretation of the uncertainty expression. Confidence levels presented on the outlook effectively informed participants' decision-making about which day to plan activities, but participants demonstrated a range of interpretations for what the term "confidence" referred to. Across all groups, participants were most likely to choose to take a walk on a day where the forecast for the area had the lowest confidence (i.e., lower likelihood of the heavy rain actually occurring). However, two-thirds of participants across all groups were not able to correctly report whether variations in confidence referred to variations in intensity, likelihood, or accuracy. When asked to describe "confidence" in their own words, there was a wide range of replies. This suggests that conveying the gist of uncertainty can be useful for decision-making, even if precise interpretation varies from the intended statistical or scientific definition. People can successfully infer meaning from qualitative descriptions of confidence, even if they are not clear on the technical meaning of the term.

For more complex decision-making tasks, clarity and transparency about uncertainty terms can help support more accurate inferences. Participants across all groups tended to report that "moderate" confidence was associated with a 50% value or above, which is greater than the 40% likelihood of the event occurring to which it was intended to refer. The findings suggest that the term "moderate" may be interpreted as a midpoint, or 50% likelihood, and this should be considered when using qualitative explanations of uncertainty. This also highlights the importance of testing the translation terms used to convert from numerical to linguistic or textual probability terms. Research has demonstrated the challenges of selecting appropriate translation thresholds, given that misinterpretation of these terms can occur due to their framing, directionality, and probabilistic format [see review in Doyle et al. (2014)]. Interpretations and translations are also influenced by biases, numeracy, and world view [see review in Doyle et al. (2020); Tversky and Kahneman 1974; Slovic et al. 2004; Joslyn et al. 2009; Kahan et al. 2012]. The inconsistent interpretations of confidence, biased perceptions of "moderate" confidence, and latent beliefs and understanding of what "confidence" means in a severe weather outlook forecast context all indicate the value in clearly defining use of the term, particularly in the context of likelihood, reliability, and consistency.

d. *Icons can be helpful, in moderation*

Weather icons may assist in mental "chunking" of information about weather processes, facilitating a high-level understanding of the distribution of severe weather in a national long-range outlook. While previous work has found that icons may not effectively describe specific weather phenomena (Zabini et al. 2015; Zabini 2016), we find that they can be helpful on maps for distinguishing patterns of distributed weather phenomena. This may be particularly useful for outlooks that show large regional areas where multiple different weather phenomena may be occurring at once. Weather icons helped participants identify that three different types of weather were occurring across the country, facilitating an accurate high-level understanding of overall severe weather patterns affecting the country. This improved comprehension and awareness could increase engagement with phenomenon-specific preparedness actions, such as seeking out information about preparing for snowfall (Griffin et al. 1999; Paton et al. 2005; Lindell and Perry 2012). However, maps that showed icons for both weather and confidence levels were not associated with increased comprehension, suggesting that too many icons can lead to confusion.

Despite the improved performance in identifying weather patterns from the outlook with the addition of weather icons, participants who viewed the map with weather icons reported a lower ease of reading information from the outlook than those who saw a map with no icons. The results suggest that weather icons can help audiences more effectively infer overall weather patterns on a map but that icons do not necessarily improve ease of reading, and too many icons can be ineffective. Experimental results for the effective use of icons in weather risk communication have been mixed, and further research on this topic could help further elaborate how to cognitively leverage the advantages of icons in such graphics.

e. *Limitations and future work*

Our targeted sample of engaged weather information users is helpful for the context of this study and application of the findings in NZ, but we acknowledge that the findings may not be transferable to all social and demographic groups. The study sample comprised highly educated long-term residents of urban NZ over the age of 40 years with a relatively high familiarity with the Severe Weather Outlook graphics. Further research with a more diverse sample could reveal deeper insights into social and cognitive processes affecting inferences and decision-making. Similarly, severe weather outlooks will have different design needs in different global and meteorological contexts and with diverse cultural audiences.

The empirical designs tested in this study are not proposed as potential solutions for use in practice. The designs have been created to carefully control for testing of visual representation variables between groups. Development of a product for practice should be designed based on the findings of this and other studies and employ user experience design and accessibility testing that incorporates end users' perspectives via the application of evidence-based usability guidelines (e.g., Tan et al. 2020). Further empirical studies of other elements of the graphic, such as the presence or absence of geographical markings, color choice, static or interactive formats, and how these large-scale products link with the higher-resolution content in weather warnings and watches, could deepen our understanding of how audiences engage with such products. It is important to note that we explore the high-level influence of outlook design on intended decision-making by comparing differences in comprehension across visual representation variables between groups, but we acknowledge that decision-making is a complex cognitive task and that other psychological biases, heuristics, and sociopolitical factors, such as trust, play into decision-making tasks. Further, we acknowledge that these findings are based on participant responses to a hypothetical scenario, and actual behavior may differ from that reported here under the stress and pressure of a real event. We note that the questions used in experiment A focused on a day in which no severe weather was forecast, and thus, the assumption is that the desired behavior is no mitigation actions need to be taken. It would be valuable to compare this with a situation in which the desired response is an increase in protective action and increased levels of concern are warranted.

Here, we tested the forecast area colors associated with the MetService control map to reduce experimental error. We suggest that future studies could explore forecast areas that are less saturated (e.g., lighter colored) or introduce other visual variables for representing forecast areas and uncertainty, such as transparency or texture. Future studies exploring the role of severe weather outlooks in decision-making could also support and inform more targeted testing about design, such as interactive maps where display variables can be adjusted by the user.

5. Conclusions

Severe weather outlooks can help audiences prepare for severe weather several days in advance and play an important role in mitigating the impacts of events such as severe gales, widespread snowfall, and heavy rain. From a study of 417 people in NZ, we find that variables in the visual representation of a severe weather outlook's time window, forecast area, icons, and uncertainty expression can affect how people make inferences and decisions regarding the outlook. On the basis of these findings, we recommend that the following points be considered when developing or designing a long-term forecast outlook product for nonspecialist audiences:

- Displaying outlook content for multiple days alongside each other in a panel graphic (e.g., Tuesday, Wednesday, Thursday), rather than collectively in a composite graphic (e.g., Tuesday–Thursday), may help support accurate inferences about the forecast over time and reduce perceived forecasting biases.
- Extrinsic expressions of uncertainty, which add objects such as text to describe uncertainty, may help support more accurate interpretations about likelihood than intrinsic expressions, which manipulate visual variables of the existing graphic, such as color saturation.
- Visual representation of the forecast area on the map can affect perceptions of confidence and likelihood. Salient solid-shaded areas may be associated with higher levels of perceived forecast confidence (in our study, higher confidence than that intended by the forecast) than areas that are displayed using outlines.
- Inclusion of weather icons in outlooks can increase comprehension of overall regional weather patterns, but adding icons for uncertainty creates visual clutter and confusion. Icons should be used in moderation and should be tested with users for ease of reading.
- Empirical testing of outlook design with relevant audiences can improve visual communication and support more informed decision-making.

Acknowledgments. The Massey University Code of Ethical of Conduct was followed for this study, and this research received a Low-Risk Ethics Notification of 4000020738 on 19 March 2019. The authors thank Peter Kreft (MetService) for his guidance and advice on this project, the participants of this study for volunteering their time and insights, Lucy Kaiser for Qualtrics support, Stephanie Raill (MetService) for providing Severe Weather Outlook web usage data, and three anonymous reviewers for their thoughtful and constructive comments on an earlier version of this paper. Author Clive was supported by the New Zealand Ministry of Business, Innovation and Employment through the Hazard and Risk Management Program (Strategic Science Investment Fund; contract C05 × 1702). Authors Clive, Doyle, and Potter were supported by the National Science Challenges: Resilience to Nature's Challenges | Kia manawaroa—Ngā Ākina o Te Ao Tūroa 2019–24. Authors Doyle and Johnston were partially supported by QuakeCoRE | Te Hiranga Rū—Aotearoa NZ Centre for Earthquake Resilience 2021, a New Zealand Tertiary Education Commission–funded center.

Data availability statement. Data can be accessed through a request to GNS Science at socialscience@gns.cri.nz.

APPENDIX

Survey Questions

Table A1 shows the survey questions used in this research. The questions explored the participants' comprehension, perceptions of likelihood, and self-reported ease of reading.

TABLE A1. An overview of the survey questions seen by participants.

No.	Question
General	
1	"How confident are you in your skills for reading and using weather forecasts?"
2	"How often do you use weather forecasts?"
3	"Overall, how long have you been visiting or residing in New Zealand?"
4	"In which of the following regions of New Zealand do you spend the most time?"
5	"The above image is an example of a MetService severe weather outlook. The outlooks show severe weather forecasted over the next 3 - 6 days. How familiar are you with these products?"
6	"How likely are you to use MetService severe weather outlooks for planning activities (e.g., work plans or family outings)?"
Expt A: time frame visualization (Fig. 2)	
7	"On which of the following days is the campsite forecasted to experience severe weather? (Tick all that apply)"
8	"How likely is it that heavy rain will occur on Friday 7 February at the campsite?"
9	"How concerned would you be to drive mountainous roads near the campsite on Friday 7 February in the forecasted conditions?"
10	"How likely would you be to change your plans to travel and camp based on this forecast?"
11	"How easy is it to read information from this map?"
Expt B: perception and expression of uncertainty (Fig. 3)	
12	"On a scale of 0 - 100%, where 0% is not at all confident and 100% is extremely confident, how confident are you that there will actually be heavy rain at Te Kaha on Friday 7 February?"
13	"On a scale of 0 - 100%, where 0% is not at all confident and 100% is extremely confident, how confident are you that there will actually be heavy rain at Te Kaha on Saturday 8 February?"
14	"Would you plan your walk for Friday, Saturday, or Sunday?"
15	"How easy is it to read information from this map?"
Expt C: icons	
16	"How many different types of weather (e.g., rain, snow, thunderstorms) are happening across the country over the outlook period?"
17	"Select all of the following that are true for areas where severe gales are forecasted:"
18	"Which part of the country is more likely to experience severe weather on Sunday?"
19	"How easy is it to read information from this map?"
Closing	
20	"How useful do you think severe weather outlooks are as a tool for planning ahead?"
21	"How likely are you to use MetService severe weather outlook for planning activities in the future (e.g., family outings or long weekend trips)?"
22	"What do you think the different levels of 'confidence' in the forecast mean?"
23	"What did you like or dislike about the maps you saw in this survey?"
24	"If you have any further comments you would like to share, please do so below:"

REFERENCES

Ash, K. D., R. L. Schumann III, and G. C. Bowser, 2014: Tornado warning trade-offs: Evaluating choices for visually communicating risk. *Wea. Climate Soc.*, **6**, 104–118, https://doi.org/10.1175/WCAS-D-13-00021.1.

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.

——, D. Cao, M. Jenkins, P. R. Pennathur, M. Farry, E. Roth, S. S. Potter, and J. Pfautz, 2011: Comparing uncertainty visualizations for a dynamic decision-making task. *J. Cognit. Eng. Decis. Making*, **5**, 277–293, https://doi.org/10.1177/1555343411415793.

Bostrom, A., L. Anselin, and J. Farris, 2008: Visualizing seismic risk and uncertainty. *Ann. N. Y. Acad. Sci.*, **1128**, 29–40, https://doi.org/10.1196/annals.1399.005.

Broad, K., A. Leiserowitz, J. Weinkle, and M. Steketee, 2007: Misinterpretations of the "cone of uncertainty" in Florida during the 2004 hurricane season. *Bull. Amer. Meteor. Soc.*, **88**, 651–668, https://doi.org/10.1175/BAMS-88-5-651.

Bryant, B., M. Holiner, R. Kroot, K. Sherman-Morris, W. Smylie, L. Stryjewski, M. Thomas, and C. Williams, 2014: Usage of color scales on radar maps. *J. Oper. Meteor.*, **2**, 169–179, https://doi.org/10.15191/nwajom.2014.0214.

Carr, R. H., B. Montz, K. Maxfield, S. Hoekstra, K. Semmens, and E. Goldman, 2016: Effectively communicating risk and uncertainty to the public: Assessing the National Weather Service's flood forecast and warning tools. *Bull. Amer. Meteor. Soc.*, **97**, 1649–1665, https://doi.org/10.1175/BAMS-D-14-00248.1.

Cavanagh, P., 2011: Visual cognition. *Vision Res.*, **51**, 1538–1551, https://doi.org/10.1016/j.visres.2011.01.015.

Cheong, L., S. Bleisch, A. Kealy, K. Tolhurst, T. Wilkening, and M. Duckham, 2016: Evaluating the impact of visualization of wildfire hazard upon decision-making under uncertainty. *Int. J. Geogr. Inf. Sci.*, **30**, 1377–1404, https://doi.org/10.1080/13658816.2015.1131829.

Clive, M. A. T., J. M. Lindsay, G. S. Leonard, C. Lutteroth, A. Bostrom, and P. Corballis, 2021: Volcanic hazard map visualisation affects cognition and crisis decision-making. *Int. J. Disaster Risk Reduct.*, **55**, 102102, https://doi.org/10.1016/j.ijdrr.2021.102102.

Creswell, J. W., 2013: *Qualitative Inquiry and Research Design: Choosing Among Five Approaches*. Sage, 448 pp.

Deitrick, S., and E. A. Wentz, 2015: Developing implicit uncertainty visualization methods motivated by theories in decision science. *Ann. Assoc. Amer. Geogr.*, **105**, 531–551, https://doi.org/10.1080/00045608.2015.1012635.

Demuth, J. L., R. E. Morss, J. K. Lazo, and D. C. Hilderbrand, 2013: Improving effectiveness of weather risk communication on the NWS point-and-click web page. *Wea. Forecasting*, **28**, 711–726, https://doi.org/10.1175/WAF-D-12-00118.1.

Dodge, M., R. Kitchin, and C. Perkins, 2011: *The Map Reader: Theories of Mapping Practice and Cartographic Representation*. John Wiley and Sons, 504 pp.

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

——, D. M. Johnston, R. Smith, and D. Paton, 2019: Communicating model uncertainty for natural hazards: A qualitative systematic thematic review. *Int. J. Disaster Risk Reduct.*, **33**, 449–476, https://doi.org/10.1016/j.ijdrr.2018.10.023.

——, J. Mc Clure, S. H. Potter, M. K. Lindell, J. S. Becker, S. A. Fraser, and D. M. Johnston, 2020: Interpretations of aftershock advice and probabilities after the 2013 Cook Strait earthquakes, Aotearoa New Zealand. *Int. J. Disaster Risk Reduct.*, **49**, 101653, https://doi.org/10.1016/j.ijdrr.2020.101653.

Egan, D. E., and B. J. Schwartz, 1979: Chunking in recall of symbolic drawings. *Mem. Cognit.*, **7**, 149–158, https://doi.org/10.3758/BF03197595.

Ernst, S., J. Ripberger, M. J. Krocak, H. Jenkins-Smith, and C. Silva, 2021: Colorful language: Investigating public interpretation of the Storm Prediction Center convective outlook. *Wea. Forecasting*, **36**, 1785–1797, https://doi.org/10.1175/WAF-D-21-0001.1.

Evans, J. St. B. T., and K. E. Stanovich, 2013: Dual-process theories of higher cognition. *Perspect. Psychol. Sci.*, **8**, 223–241, https://doi.org/10.1177/1745691612460685.

Fabrikant, S. I., S. R. Hespanha, and M. Hegarty, 2010: Cognitively inspired and perceptually salient graphic displays for efficient spatial inference making. *Ann. Assoc. Amer. Geogr.*, **100**, 13–29, https://doi.org/10.1080/00045600903362378.

Gill, J., and Coauthors, 2008: Guidelines on communicating forecast uncertainty. WMO/TD-4122, 25 pp., https://library.wmo.int/doc_num.php?explnum_id=4687.

Gould, S. F., N. J. Beeton, R. M. B. Harris, M. F. Hutchinson, A. M. Lechner, L. L. Porfirio, and B. G. Mackey, 2014: A tool for simulating and communicating uncertainty when modelling species distributions under future climates. *Ecol. Evol.*, **4**, 4798–4811, https://doi.org/10.1002/ece3.1319.

Grant, E. R., and M. J. Spivey, 2003: Eye movements and problem solving. *Psychol. Sci.*, **14**, 462–466, https://doi.org/10.1111/1467-9280.02454.

Griffin, R. J., S. Dunwoody, and K. Neuwirth, 1999: Proposed model of the relationship of risk information seeking and processing to the development of preventive behaviors. *Environ. Res.*, **80**, S230–S245, https://doi.org/10.1006/enrs.1998.3940.

Harold, J., I. Lorenzoni, T. F. Shipley, and K. R. Coventry, 2016: Cognitive and psychological science insights to improve climate change data visualization. *Nat. Climate Change*, **6**, 1080–1089, https://doi.org/10.1038/nclimate3162.

Hegarty, M., 2011: The cognitive science of visual–spatial displays: Implications for design. *Top. Cognit. Sci.*, **3**, 446–474, https://doi.org/10.1111/j.1756-8765.2011.01150.x.

——, M. S. Canham, and S. I. Fabrikant, 2010: Thinking about the weather: How display salience and knowledge affect performance in a graphic inference task. *J. Exp. Psychol. Learn. Mem. Cognit.*, **36**, 37–53, https://doi.org/10.1037/a0017683.

Hudson-Doyle, E. E., D. Paton, and D. M. Johnston, 2018: Reflections on the communication of uncertainty: Developing decision-relevant information. *Proc. ISCRAM Asia Pacific Conf. 2018*, Wellington, New Zealand, International Association for Information Systems for Crisis Response and Management, 24 pp., https://idl.iscram.org/files/emmahudson-doyle/2018/1650_EmmaHudson-Doyle_etal2018.pdf.

Itti, L., and C. Koch, 2000: A saliency-based search mechanism for overt and covert shifts of visual attention. *Vision Res.*, **40**, 1489–1506, https://doi.org/10.1016/s0042-6989(99)00163-7.

Jarvenpaa, S. L., 1990: Graphic displays in decision making—The visual salience effect. *J. Behav. Decis. Making*, **3**, 247–262, https://doi.org/10.1002/bdm.3960030403.

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

Joslyn, S. L., 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.

Kahan, D. M., E. Peters, M. Wittlin, P. Slovic, L. L. Ouellette, D. Braman, and G. Mandel, 2012: The polarizing impact of science literacy and numeracy on perceived climate change risks. *Nat. Climate Change*, **2**, 732–735, https://doi.org/10.1038/nclimate1547.

Kinkeldey, C., A. M. MacEachren, and J. Schiewe, 2014: How to assess visual communication of uncertainty? A systematic review of geospatial uncertainty visualisation user studies. *Cartographic J.*, **51**, 372–386, https://doi.org/10.1179/1743277414Y.0000000099.

Kolbeinsson, A., G. Falkman, and J. Lindblom, 2015: Showing uncertainty in aircraft cockpits using icons. *Procedia Manuf.*, **3**, 2905–2912, https://doi.org/10.1016/j.promfg.2015.07.805.

Kubíček, P., and Č. Šašinka, 2011: Thematic uncertainty visualization usability – Comparison of basic methods. *Ann. GIS*, **17**, 253–263, https://doi.org/10.1080/19475683.2011.625978.

Kunz, M., A. Grêt-Regamey, and L. Hurni, 2011: Visualization of uncertainty in natural hazards assessments using an interactive cartographic information system. *Nat. Hazards*, **59**, 1735–1751, https://doi.org/10.1007/s11069-011-9864-y.

Lazo, J. K., R. E. Morss, and J. L. Demuth, 2009: 300 billion served. *Bull. Amer. Meteor. Soc.*, **90**, 785–798, https://doi.org/10.1175/2008BAMS2604.1.

Lindell, M. K., and R. W. Perry, 2012: The protective action decision model: Theoretical modifications and additional evidence. *Risk Anal.*, **32**, 616–632, https://doi.org/10.1111/j.1539-6924.2011.01647.x.

MacEachren, A. M., 1992: Visualizing uncertain information. *Cartogr. Perspect.*, **13**, 10–19, https://doi.org/10.14714/CP13.1000.

——, 1995: *How Maps Work: Representation, Visualization, and Design*. Guilford Press, 526 pp.

Mastrandrea, M. D., and Coauthors, 2010: Guidance note for lead authors of the IPCC fifth assessment report on consistent treatment of uncertainties. IPCC Guidance Note, 7 pp., ipcc.ch/site/assets/uploads/2018/05/uncertainty-guidance-note.pdf.

Miller, G. A., 1956: The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychol. Rev.*, **63**, 81–97, https://doi.org/10.1037/h0043158.

Morss, R. E., J. L. Demuth, and J. K. Lazo, 2008: Communicating uncertainty in weather forecasts: A survey of the U.S. public. *Wea. Forecasting*, **23**, 974–991, https://doi.org/10.1175/2008WAF2007088.1.

NEMA, 2015: The guide to the National Civil Defence emergency management plan. Department of the Prime Minister and Cabinet, 321 pp., https://www.civildefence.govt.nz/assets/guide-to-the-national-cdem-plan/Guide-to-the-National-CDEM-Plan-2015.pdf.

New Zealand Mountain Safety Council, 2018: *A Walk in the Park? A Deep Dive into Tramping Incidents in New Zealand*. New Zealand Mountain Safety Council, 94 pp.

Padilla, L. M., S. H. Creem-Regehr, M. Hegarty, and J. K. Jeanine, 2018: Decision making with visualizations: A cognitive framework across disciplines. *Cognit. Res.*, **3**, 29, https://doi.org/10.1186/s41235-018-0120-9.

Pang, A., 2008: Visualizing uncertainty in natural hazards. *Risk Assessment, Modeling and Decision Support, Strategic Directions*, A. Bostrom, S. French, and S. Gottlieb, Eds., Risk, Governance and Society, Vol. 14, Springer, 261–294.

Paton, D., L. Smith, and D. Johston, 2005: When good intentions turn bad: Promoting natural hazard preparedness. *Aust. J. Emerg. Manage.*, **20**, 25–30.

Patt, A., and S. Dessai, 2005: Communicating uncertainty: Lessons learned and suggestions for climate change assessment. *C. R. Geosci.*, **337**, 425–441, https://doi.org/10.1016/j.crte.2004.10.004.

Patterson, R. E., L. M. Blaha, G. G. Grinstein, K. K. Liggett, D. E. Kaveney, K. C. Sheldon, P. R. Havig, and J. A. Moore, 2014: A human cognition framework for information visualisation. *Comput. Graphics*, **42**, 42–58, https://doi.org/10.1016/j.cag.2014.03.002.

Perceptive, 2017: MetService severe weather reporting. Market Research Rep., 25 pp.

Potter, S. H., P. V. Kreft, P. Milojev, C. Noble, B. Montz, A. Dhellemmes, R. J. Woods, and S. Gauden-Ing, 2018: The influence of impact-based severe weather warnings on risk perceptions and intended protective actions. *Int. J. Disaster Risk Reduct.*, **30**, 34–43, https://doi.org/10.1016/j.ijdrr.2018.03.031.

Reed, J. R., and J. C. Senkbeil, 2020: Perception and comprehension of the extended forecast graphic: A survey of broadcast meteorologists and the public. *Bull. Amer. Meteor. Soc.*, **101**, E221–E236, https://doi.org/10.1175/BAMS-D-19-0078.1.

——, S. A. Jasko, and J. C. Senkbeil, 2022: Weather forecast semiotics: Public interpretation of common weather icons. *Wea. Climate Soc.*, **14**, 323–335, https://doi.org/10.1175/WCAS-D-21-0067.1.

Ripberger, J. T., H. C. Jenkins-Smith, C. L. Silva, D. E. Carlson, and M. Henderson, 2014: Social media and severe weather: Do tweets provide a valid indicator of public attention to severe weather risk communication? *Wea. Climate Soc.*, **6**, 520–530, https://doi.org/10.1175/WCAS-D-13-00028.1.

Robinson, A. H., 1967: Psychological aspects of color in cartography. *Int. Yearb. Cartogr.*, **7**, 50–59.

Roth, R. E., and Coauthors, 2017: User studies in cartography: opportunities for empirical research on interactive maps and visualizations. *Int. J. Cartogr.*, **3**, 61–89, https://doi.org/10.1080/23729333.2017.1288534.

Severtson, D. J., and C. Vatovec, 2012: The theory-based influence of map features on risk beliefs: Self-reports of what is seen and understood for maps depicting an environmental health hazard. *J. Health Commun.*, **17**, 836–856, https://doi.org/10.1080/10810730.2011.650933.

——, and J. D. Myers, 2013: The influence of uncertain map features on risk beliefs and perceived ambiguity for maps of modeled cancer risk from air pollution. *Risk Anal.*, **33**, 818–837, https://doi.org/10.1111/j.1539-6924.2012.01893.x.

Sivle, A. D., S. D. Kolstø, P. J. K. Hansen, and J. Kristiansen, 2014: How do laypeople evaluate the degree of certainty in a weather report? A case study of the use of the web service yr.no. *Wea. Climate Soc.*, **6**, 399–412, https://doi.org/10.1175/WCAS-D-12-00054.1.

Slocum, T. A., D. C. Cliburn, J. J. Feddema, and J. R. Miller, 2003: Evaluating the usability of a tool for visualizing the uncertainty of the future global water balance. *Cartogr. Geogr. Inf. Sci.*, **30**, 299–317, https://doi.org/10.1559/152304003322606210.

Slovic, P., M. L. Finucane, E. Peters, and D. G. MacGregor, 2004: Risk as analysis and risk as feelings: Some thoughts about affect, reason, risk, and rationality. *Risk Anal.*, **24**, 311–322, https://doi.org/10.1111/j.0272-4332.2004.00433.x.

Spiegelhalter, D., M. Pearson, and I. Short, 2011: Visualizing uncertainty about the future. *Science*, **333**, 1393–1400, https://doi.org/10.1126/science.1191181.

Tan, M. L., R. Prasanna, K. Stock, E. E. H. Doyle, G. Leonard, and D. Johnston, 2020: Understanding end-users' perspectives: Towards developing usability guidelines for disaster apps. *Prog. Disaster Sci.*, **7**, 100118, https://doi.org/10.1016/j.pdisas.2020.100118.

Thalmann, M., A. S. Souza, and K. Oberauer, 2019: How does chunking help working memory? *J. Exp. Psychol. Learn. Mem. Cognit.*, **45**, 37–55, https://doi.org/10.1037/xlm0000578.

Thompson, M. A., J. M. Lindsay, and J. C. Gaillard, 2015: The influence of probabilistic volcanic hazard map properties on hazard communication. *J. Appl. Volcanol.*, **4**, 6, https://doi.org/10.1186/s13617-015-0023-0.

Tversky, A., and D. Kahneman, 1974: Judgment under uncertainty: Heuristics and biases. *Science*, **185**, 1124–1131, https://doi.org/10.1126/science.185.4157.1124.

Visschers, V. H. M., R. M. Meertens, W. W. F. Passchier, and N. N. K. de Vries, 2009: Probability information in risk communication: A review of the research literature. *Risk Anal.*, **29**, 267–287, https://doi.org/10.1111/j.1539-6924.2008.01137.x.

Wallsten, T. S., S. Fillenbaum, and J. A. Cox, 1986: Base rate effects on the interpretations of probability and frequency expressions. *J. Mem. Lang.*, **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. *J. Exp. Psychol.*, **16**, 781–789, https://doi.org/10.1037/0096-1523.16.4.781.

Wright, K. C., D. M. Johnston, and J. S. Becker, 2010: The recognition and understanding of severe weather advice in New Zealand; results of a 2009 survey. GNS Science Rep., 11 pp., https://eurekamag.com/research/037/184/037184360.php.

Zabini, F., 2016: Mobile weather apps or the illusion of certainty. *Meteor. Appl.*, **23**, 663–670, https://doi.org/10.1002/met.1589.

——, V. Grasso, R. Magno, F. Meneguzzo, and B. Gozzini, 2015: Communication and interpretation of regional weather forecasts: A survey of the Italian public. *Meteor. Appl.*, **22**, 495–504, https://doi.org/10.1002/met.1480.