Improving Effectiveness of Weather Risk Communication on the NWS Point-and-Click Web Page

JULIE L. DEMUTH, REBECCA E. MORSS, AND JEFFREY K. LAZO

National Center for Atmospheric Research,* Boulder, Colorado

DOUGLAS C. HILDERBRAND

Meteorological Developmental Laboratory, NOAA/National Weather Service/Office of Science and Technology, Silver Spring, Maryland

(Manuscript received 4 November 2012, in final form 28 February 2013)

ABSTRACT

The National Weather Service’s (NWS) point-and-click (PnC) web page is a primary channel through which NWS directly provides routine and hazardous weather information to its users. The research presented here aims to improve risk communication of hazardous weather information on the PnC web page. The focus is on improving communication of threat existence and threat timing because this important information influences how individuals perceive and respond to a weather risk. Experimental presentations of PnC forecast information were designed for two weather scenarios: a severe thunderstorm warning and a flood watch. The experimental presentations were created by adding new textual and graphical pieces of information that were intended to better convey threat existence and timing, and they were evaluated through two rounds of nationwide surveys of PnC web page users. The survey results show that the default presentation of forecast information on the PnC web page was the least effective at conveying hazardous weather threat existence and timing. Adding start-time text and end-time text, when these information pieces were coupled, helped respondents understand the precise time that weather threats were in effect for the rapid-onset, short-duration severe thunderstorm warning and for the delayed-start, longer-duration flood watch. Adding a box graphic placed around the forecast icons further enhanced communication effectiveness by drawing respondents’ attention to the weather threat. Other experimental forecast presentations were designed but were less effective at communicating hazardous weather threat existence and timing, illustrating the importance of empirically evaluating weather risk communication prior to providing it operationally.

1. Introduction

The National Weather Service’s (NWS’s) point-and-click (PnC) web page (Fig. 1) is a primary channel through which NWS directly provides routine and hazardous weather information to its users. Recognizing the importance of the PnC web page, the NWS funded an applied research effort aimed at characterizing the key strengths and limitations of the PnC information and at identifying potential areas for improvement1 [see Demuth et al. (2012a) for details of the full research effort]. The PnC web page was found to be valued by its users for many reasons, chief among them that it provides accurate, up-to-date forecast information for a specific geographic area in usable
formats both at a glance (i.e., the forecast icons) and in greater detail.\footnote{The PnC forecast information is populated by the Graphical Forecast Editor (GFE) forecast grids that are updated at least twice daily by forecasters in each NWS Weather Forecast Office. Currently, there is a unique PnC forecast for every 2.5-km$^2$ grid in the continental United States, Hawaii, and Guam; for every 1.25-km$^2$ grid in Puerto Rico; and at least for every 6-km$^2$ grid in Alaska (NOAA 2013).}

However, a key limitation that was identified in the initial phase of the research effort is that several aspects of hazardous weather risk information are not effectively communicated on the PnC web page. These aspects include the existence of a hazardous weather threat, a threat’s timing, a threat’s location and spatial extent, and the importance of a threat relative to other ones denoted on the PnC. Effectively communicating this information is important because it can help PnC users make potentially life-saving decisions when hazardous weather threatens. To address this shortcoming, the subsequent phase of the research effort, described here, focused on ways to improve how the PnC web page communicates two of the hazardous weather risk aspects—threat existence and threat timing.

We focused on more effectively communicating whether a hazardous weather threat exists and for what time period because these are two fundamental pieces of information required for individuals to begin mapping a potential weather risk onto their lives and then determining if and how to respond. At the most basic level, effective communication requires that information be received (via some sensory input, e.g., audibly or visually) and understood (Renn 2008). Identifying the existence of a weather threat via a web page, such as the PnC, requires the user to attend to and notice the information. This is influenced by multiple factors, including the individual’s ability and motivation (Renn 2008; Hawkins and Daly 1988) and information-processing capacity (Lang 2000; Mayer 2005), as well as the visual accessibility and distinctiveness of the information (Mayer 2005; Hillstrom and Chai 2006). The existence of all hazardous weather threats is conveyed on the PnC web page only as red, underlined text (which is a hyperlink to the weather product text details) located under the “Hazardous weather conditions” headline\footnote{Hazardous weather was denoted in this way on the PnC web page at the time of the study and also on the current redesigned web page.} (e.g., see the “flood watch” text in Fig. 1). Because hazardous weather events occur over finite periods of time, understanding a threat’s timing is also important for evaluating weather risks. Unlike some types of

![Fig. 1. Example PnC forecast. This real forecast served as the baseline for developing the experimental forecasts. Layout and content shown are how the PnC forecast web page was presented for over a decade up until 2 Jul 2012, when NWS fielded a redesigned PnC forecast web page.](image-url)
specific times. Yet, the timing of hazardous weather threats is not explicitly conveyed on the PnC web page (Fig. 1).

Working closely with the NWS, we developed experimental graphical and textual pieces of information aimed at more effectively conveying hazardous weather threat existence and timing. We created experimental presentations of the PnC forecasts by adding the information pieces for two hazardous weather scenarios: a severe thunderstorm warning and a flood watch. The experimental presentations were evaluated in two rounds of nationwide surveys of PnC web page users that asked questions about respondents’ 1) identification of weather threat existence and understanding of threat timing, 2) perceptions of how well the forecasts conveyed the threat existence and timing and of the forecasts’ visual aesthetics, and 3) preferences for the different pieces of graphical and textual information that were added.

Section 2 describes our research design, the added graphical and textual information evaluated in each survey, and the survey implementation and sample characteristics. The results of the first and second surveys are reported upon in sections 3 and 4, respectively. In section 5, we summarize the findings and discuss their broader implications.

2. Research design

The full research effort assessed multiple aspects of the PnC web page in several phases. Prior to the research phase discussed here, we conducted focus groups, a usability evaluation (Zimmerman et al. 2010), and a survey that, collectively, identified a variety of strengths and limitations of the PnC web page discussed in detail in Demuth et al. (2012a). Among these initial findings were that getting information about potential hazardous weather threats is important to PnC web page users, but that hazardous weather threats are not prominently or clearly indicated, nor is their timing clearly conveyed. This earlier work motivated the research discussed in this paper, aimed at examining ways to more effectively communicate threat existence and timing on the PnC.

This section describes the research design for this portion of the research effort, including the weather scenarios we examined, our survey design and questions, the experimental forecast presentations we designed and evaluated, and the survey implementation, our sampling approach, and sample characteristics. Further details are available in Demuth et al. (2012a).

a. Weather scenarios and survey design

The first step in the research design was selecting weather scenarios to examine. NWS issues routine forecasts for 12-h time periods—from 0600 to 1800 and from 1800 to 0600 local time (LT)—out to 7 days. In the PnC forecast shown in Fig. 1, for example, the first forecast period, “late afternoon,” spans from the current time through 1800 local time; the second forecast period, “tonight,” spans from 1800 to 0600 LT the next day; and so forth. NWS also issues event-driven forecast products (e.g., watches, warnings, advisories) when hazardous weather threatens. Of the many possible scenarios of weather threat type, products, and timing, we chose two for our study: 1) a severe thunderstorm warning that begins and ends in the first forecast period (i.e., a rapid-onset, short-duration threat) and 2) a flood watch that begins and ends in later forecast periods (i.e., a delayed-start, longer-duration threat).

Next we developed experimental forecast presentations of the weather scenarios portrayed on the PnC web page. From the NWS’s web page we captured a real PnC image with forecast information (i.e., icons, text, radar, satellite) that could realistically represent both the severe thunderstorm warning and flood watch scenarios (Fig. 1). This image served as the baseline for developing all of the experimental presentations. The geographic referents (city, state, latitude, longitude, elevation) were modified so that the real location would not affect study participants’ assessment of the forecast4 (see Figs. 2 and 3). We created the experimental presentations by adding new graphical and textual pieces of information (hereafter referred to as attributes) that were intended to improve communication of the threat existence and timing. We drew from the information design (Redish 2000) and web site usability testing (DHHS 2006; Zimmerman and Akerelrea 2003) literature to design the attributes, described in the following subsections. We also collaborated with NWS web-technology personnel to ensure that the attributes could be operationally generated and displayed on the PnC web page without requiring excessive computing power and bandwidth. Other than the added attributes, all other information (e.g., icons, forecast text, date, time, current conditions) remained constant across the experimental presentations. The forecast presentation without any added attributes served as the control, representing the way the PnC forecast was provided on the NWS web page at the time of the study.

4This is one way we attempted to minimize threats to internal validity. Internal validity is the extent to which a causal relationship exists between independent and dependent variables. Internal validity of a study is threatened when an extraneous variable—for example, in this case, the real forecast location of Billings, Montana—unintentionally influences study participants, confounding the effects of the independent variable (Shadish et al. 2001).
The target population for this study is all users for all NWS PnC web pages. Working with the NWS, we recruited 88,191 individuals via the PnC web page who were willing to participate in our study (Demuth et al. 2012a), and we randomly selected individuals from this list to participate in one of two controlled-access web surveys (see section 2d). In both surveys, we first evaluated the experimental presentations between subjects, where participants were randomly assigned to one of the experimental presentations or to the control forecast and then were asked a series of questions about the image they received. The questions assessed respondents’ ability to identify the severe thunderstorm warning or flood watch threat in the forecast, ability to accurately identify the threat start and end times, and perceptions of the forecast information. The experimental forecast presentation was shown with each question so that respondents could refer to it. We then evaluated the experimental presentations within subjects, where respondents were shown a series of paired experimental forecast presentations. Each pair included an experimental presentation that had a single attribute.
alongside the forecast without that attribute (i.e., the control), and respondents were asked which they preferred. All respondents were asked their preferences about each attribute, and they received the questions in the same order. Additional survey questions asked about respondents’ experience using the PnP web page, self-reported knowledge of the PnP web page, and demographic characteristics.

We did an initial test of the attributes in a first round of surveys (hereafter survey 1). Based on the results, we refined the attribute set and implemented a second round of surveys (hereafter survey 2). Each round included a severe thunderstorm warning scenario survey and a parallel flood watch scenario survey with respective sets of experimental forecasts. The specific attributes we developed and evaluated are described in the following two sections.

b. Experimental forecast presentations for survey 1

For survey 1, we created and evaluated three attributes. One was a bar placed underneath the forecast-at-a-glance icons. The bar position and length were
intended to represent the time period over which the threat was in effect. The bar was colored red for warnings and orange for watches, and it was labeled with text identifying the threat (e.g., flood watch). We added the bar to both the severe thunderstorm warning and flood watch scenario experimental forecasts.

A second attribute was end-time text added to the red, underlined text that identifies hazardous weather on the PnC web page. The end-time text was intended to indicate when the threat expires. We initially included the end-time text primarily with rapid-onset, short-duration warnings in mind. Such warnings are in effect the moment a NWS forecaster issues them, but their expiration time is not conveyed on the PnC web page. Instead, a user must click on and read the hazardous weather product to glean this information. We added the end-time text to both the severe thunderstorm warning and flood watch scenario experimental forecasts.

A third attribute was added for the severe thunderstorm warning scenario only: a box placed around the forecast-at-a-glance icon in the first forecast period. The box was intended to draw attention to the severe thunderstorm warning, which is a type of rapid-onset, short-duration weather hazard that can pose imminent threats to life and property. It was colored red and labeled with text identifying the threat.

We created the experimental forecasts by adding each of the single attributes (e.g., bar only) and by adding all possible combinations of attributes (e.g., bar + end-time text + box). Thus, there were four distinct experimental forecasts for the flood watch scenario and eight for the severe thunderstorm warning scenario, including the control forecasts, which had no added attributes. Figure 2 shows two example experimental forecasts for each scenario: one with a single attribute and one with all attributes combined.

c. Experimental forecast presentations for survey 2

As we will discuss in section 3, there were mixed results regarding the effectiveness of the attributes evaluated in survey 1 for communicating the threat existence and timing. Therefore, we refined the attribute set based on the survey 1 results and developed and evaluated a revised set of experimental forecast presentations in survey 2.

We retained the end-time text and added the start-time text attribute to the red, underlined text that identifies hazardous weather on the PnC web page. The start-time text was intended to indicate when a hazardous weather threat goes into effect or if one already is in effect. We also retained the box attribute, and in addition to using it in the severe thunderstorm warning scenario, we added it to the flood watch scenario by placing it around the series of forecast-at-a-glance icons during which the threat was forecast. The watch box was colored orange (see footnote 5) and labeled with text identifying the threat. We omitted the bar attribute from the survey 2 experimental forecast presentations.

All three attributes were used in the severe thunderstorm warning and flood watch scenarios. We again created experimental forecasts [using the same baseline forecast (Fig. 1) as was used in survey 1] with each of the single attributes and with all possible combinations of attributes, so there were eight distinct experimental forecasts for each scenario, including the control forecasts with no attributes. Figure 3 shows two example experimental forecasts for each scenario: one with a single attribute and one with all attributes combined.

d. Survey implementation and sample characteristics

We developed surveys 1 and 2 and had them reviewed for scientific and technical accuracy by a select group of NWS staff. A survey research company, ResearchExec, programmed and hosted the web survey. We pretested the survey in person with individuals randomly selected from the list of recruited PnC users who reside in the Boulder, Colorado, area. The in-person pretests helped us evaluate the web survey functionality as well as respondents’ understanding of and interest in the survey questions and their completion time. We refined and finalized the surveys based on the expert and public pretests.

The implementation and sample statistics for the surveys are summarized in Table 1. As discussed in section 2a, the target population for this study is NWS PnC web page users. No comprehensive list of the PnC user population exists from which to sample [a common sampling problem when studying web page users; e.g., Couper (2000); Van Selm and Jankowski (2006)]. Thus, to sample NWS PnC users, we recruited PnC users through a link on the PnC page as described in Demuth

5 The NWS has prototyped and obtained public feedback on a new web-based watch–warning–advisory map where warnings are denoted in red, watches in orange, and advisories in yellow (Scharfenberg et al. 2012). This color scheme is commensurate with typical color associations of U.S. adults (Sue and Ritter 2007); therefore, we emulated it in the graphical attributes added to the experimental forecast presentations.

6 The NWS reviewers were selected to represent a range of expertise and perspectives regarding the PnC web page. They included NWS Headquarters employees who manage the web-based forecast policies, NWS technical staff who manage the PnC web page programming, and NWS forecasters whose forecasts populate the PnC.
et al. (2012a, appendix B) and randomly selected individuals from this list to invite to participate in the surveys. ResearchExec managed all e-mail invitations and data collection. Each invitee was provided with a unique survey web link and could only respond one time. The survey response rates, which ranged from 40.0% to 45.6%, are higher than average for Internet-based surveys (Smyth and Pearson 2011). The final sample sizes for our data analysis ranged from N = 2081 to 4239. The median survey response times ranged from 23 to 24 min.

The survey samples are nationwide with respondents from every U.S. state, indicating geographic diversity in our samples. Survey respondents were not required to answer questions about their demographic characteristics. Respondents who did provide these data were predominantly older (median 54–55 yr), Caucasian (92%–93%), and male (72%–73%). Respondents tended to be well educated (median having completed a bachelor’s degree) with a median income ranging from $75,000 to $99,999, and a majority are employed full time (57%–59%). Respondents also tended to have resided within 50 mi of their current location for a substantial period of time (median 21–22 yr) and have a median of two people per household. Our sampling approach was aimed at gathering data from individuals who are as representative as possible of the NWS PnC user population. Although this sample is likely more representative of the target population than the uncontrolled web surveys that are often conducted of weather information users, it is not designed to be representative of non-NWS web page users (cf. Lazo et al. 2009). Future work will compare the demographics of these NWS PnC web page users with those of the “general” U.S. public (i.e., non-PnC users).

Measures of respondents’ experience using the PnC web page reveal that most are long-term, frequent users of it. The vast majority (86%–90%) has been using the PnC web page for more than 3 yr. Most respondents (76%–81%) access the PnC web page at least once daily, and approximately half (49%–53%) access the PnC at least twice daily. Respondents tend to use the PnC web page quickly to glean their desired information, with most (66%–70%) spending less than 3 min on the page during a typical visit. Finally, the majority of participants (62%–71%) consider themselves “very” or “extremely” knowledgeable about the PnC web page overall, and less than 5% indicated they were “not at all” or only “a little” knowledgeable.

3. Results of survey 1

This section presents the survey 1 results in which we evaluated the bar, end-time text, and box (in the severe thunderstorm warning scenario only) attributes. We focus on presenting the most interesting and relevant results for informing survey 2; full results for every experimental forecast presentation can be found in Demuth et al. (2012a). The results discussed below are the same in each weather scenario unless otherwise indicated.

a. Threat identification and timing

In the between-subjects portion of the survey, respondents were randomly assigned to one of the experimental forecast presentations or to the control and asked questions about the image they received. To

### Table 1. Survey implementation and sample statistics.

<table>
<thead>
<tr>
<th>Study</th>
<th>No. of randomly sampled participants invited via e-mail</th>
<th>No. of e-mail bounces</th>
<th>No. of completed survey responses</th>
<th>Survey response rate (%)</th>
<th>Final sample size for data analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>10 000</td>
<td>442</td>
<td>4358</td>
<td>45.6</td>
<td>4239</td>
</tr>
<tr>
<td>1 flood watch</td>
<td>5000&lt;sup&gt;b&lt;/sup&gt;</td>
<td>223</td>
<td>2118</td>
<td>44.3</td>
<td>2081</td>
</tr>
<tr>
<td>2</td>
<td>10 000</td>
<td>578</td>
<td>3766</td>
<td>40.0</td>
<td>3717</td>
</tr>
<tr>
<td>2 flood watch</td>
<td>10 000</td>
<td>579</td>
<td>3795</td>
<td>40.3</td>
<td>3747</td>
</tr>
</tbody>
</table>

<sup>a</sup> The individuals who were invited to participate in survey 1 were removed from the sampling frame before sampling for survey 2. 

<sup>b</sup> As discussed in sections 2b and 2c, the flood watch scenario in survey 1 consisted of only four experimental forecast presentations vs the eight presentations that each other survey scenario comprised; thus, we invited half as many people to participate in the survey 1 flood watch scenario.

<sup>c</sup> Although our sampling frame consists of individuals who volunteered their e-mail address via the PnC web page, e-mail can be shared by multiple people in a household. Therefore, we included two survey questions to ensure that each respondent 1) knows of the NWS and 2) is a user of the PnC web page. The final sample sizes for data analysis omit respondents who indicated “no” to either question.

---

<sup>7</sup> Ranges are across the survey samples.
assess how well the experimental presentations helped respondents identify a threat, respondents were asked which type of hazardous weather appeared in the forecast and were offered five response options: the correct weather threat ("severe thunderstorm warning" and "flood watch" for the respective weather scenarios), three incorrect weather threats, and "I don’t know" [see Demuth et al. (2012a) for detailed response options for all survey questions discussed]. As shown in Fig. 4, regardless of the experimental forecast, the vast majority of respondents correctly identified the threat. This suggests that all of the forecast presentations were effective at communicating the existence of a threat, at least when respondents were explicitly asked to examine the forecast they were given. It is unknown to what extent respondents would have attended to the weather threats if they were not prompted.

To assess how well the experimental forecast presentations conveyed threat timing, respondents were then asked questions about what time the severe thunderstorm warning or flood watch starts and then what time it ends. Five response options were offered for each timing question: three options with different times, "I cannot tell," and "Other". For the severe thunderstorm warning scenario, the correct start and end times were "it has already started" and "6 PM on Wednesday." For the flood watch scenario, the correct start and end times were "Sometime on Thursday" and "6 PM on Saturday."

The control forecasts, which represent how the PnC forecast information was presented at the time of the study and do not provide explicit timing information about a hazardous weather threat, were the least effective at conveying threat timing (Fig. 5). Respondents

---

8 As part of our data analysis, we conducted chi-squared ($\chi^2$) tests of independence statistical tests. Used for categorical data, they test whether the conditional distributions of the dependent variable (e.g., threat identification) are identical across the experimental forecast presentations in each weather scenario. Independence means that the probability of any particular response is the same regardless of the experimental presentation. If the conditional distributions vary by forecast, they are dependent; that is, they differ by experimental presentation (Agresti and Finlay 2009). We adopted a stringent $p \leq 0.01$ level when interpreting the $\chi^2$ statistical results to minimize the chance of incorrectly concluding that there are differences among the experimental presentations for a given dependent variable (i.e., a type I error). With each $\chi^2$ statistical result, we also report the effect size (as Cramer’s $V$), which is a measure of the strength of association between the independent variable (i.e., experimental presentations) and dependent variable. Values of Cramer’s $V$ range from 0 to 1, with larger values representing stronger associations (Hayes 2005).

---

9 Severe thunderstorm warning scenario start-time options were “It has already started,” “6 PM on Wednesday,” and “Sometime tonight”; end-time options were “It has already ended,” “6 PM on Wednesday,” and “Sometime tonight.” Flood watch scenario start-time options were “It has already started,” “Sometime Thursday,” and “6 PM on Saturday”; end-time options were “It has already ended,” “Sometime Thursday,” and “6 PM on Saturday.”
who received the control were least likely to accurately identify the threat start or end time in most cases. The control also was associated with the most responses of “I cannot tell” (not shown). Because our goal was to improve people’s understanding of threat timing, here (and in survey 2) we focus on people being able to identify the correct start and end time rather than responding “I cannot tell.” These results support the finding from the initial phase of the research project that the PnC web page does not communicate important hazardous weather threat timing information well.

The experimental forecast presentation with the bar was slightly better than the control at conveying the threat timing, but it was not as effective as expected, with less than half of respondents accurately identifying either the start or end time (Fig. 5). We anticipated that the bar would be especially useful for conveying the timing of the flood watch threat because it begins in a later forecast period and spans multiple periods, but less than 34% of respondents accurately identified the start or end time in that scenario. Some respondents indicated at the end of the survey that they did not initially understand that the bar represented threat timing but that they learned this as the survey progressed. Although PnC users could develop understanding of the bar meaning through repeated exposure, the results nevertheless suggest that this attribute is not intuitively understandable to them.

Of respondents who received experimental presentations with the end-time text, not surprisingly, the vast majority accurately identified what time the hazardous weather threat expires (Fig. 5). This result holds for the forecast presentations in which the end-time text is combined with the other attributes (not shown). Surprisingly, however, in the flood watch scenario, respondents who received the end-time text forecast struggled to understand the threat start time, with less than 5% accurately identifying it. Instead, 75% of respondents who received that forecast erroneously believed that the flood watch threat had already begun; this is significantly greater than the 29% of respondents who received the control forecast who also believed the threat had already begun (not shown). These results suggest that receiving only the end-time information may have misled many respondents into thinking that the weather threat was currently in effect. A similar effect occurred in the severe thunderstorm warning scenario: 82% of respondents who received the end-time text forecast believed that the threat had already begun, compared to only 34% of those who received the control. In the severe thunderstorm scenario, however, this was the correct start time.

Most respondents who received the forecast presentation with the box accurately identified the start time (Fig. 5). However, only 7% accurately identified the end time, and more than 60% indicated they could not tell when the threat ends (not shown). This suggests that the box, on its own, is not very effective at communicating threat timing. However, the survey also included questions asking respondents’ about their perceptions of the experimental forecast presentations. On a scale from 1 (not at all) to 5 (extremely), respondents who received the box forecast rated it significantly more attention getting (mean = 4.21) than either the control or any other single-attribute forecast presentation, \( F(\text{df} = 3, \ n = 2107) = 13.37, \ p < 0.001 \). This suggests that the box

---

**FIG. 5.** Percentage of survey 1 respondents who identified the correct start and end times by experimental forecast for the (a) severe thunderstorm warning scenario [start time: \( \chi^2(\text{df} = 7, \ n = 4239) = 814.24, \ p < 0.001, \text{Cramer’s } V = 0.44 \); end time: \( \chi^2(\text{df} = 7, \ n = 4239) = 3236.09, \ p < 0.001, \text{Cramer’s } V = 0.87 \)] and (b) flood watch scenario [start time: \( \chi^2(\text{df} = 3, \ n = 2081) = 206.81, \ p < 0.001, \text{Cramer’s } V = 0.32 \); end time: \( \chi^2(\text{df} = 3, \ n = 2081) = 1434.08, \ p < 0.001, \text{Cramer’s } V = 0.83 \).
may offer benefits beyond conveying timing information; we explore this further in survey 2.

b. Attribute preferences

Next, in the within-subjects portion of the survey, respondents were shown a series of paired experimental forecast presentations—each pair consisting of a single-attribute forecast presentation alongside the forecast presentation without that attribute (i.e., the control). For each pair, respondents were asked to indicate which forecast they preferred or whether they had “no opinion/no preference.” Approximately two-thirds of respondents indicated they prefer the forecast presentation with the bar, yet about one-fifth of respondents indicated they would rather not have a forecast with this attribute (Fig. 6). Additional analysis revealed that respondents who received and repeatedly evaluated an experimental forecast that included the bar attribute (during the between-subjects part of the survey) were no more likely than those who did not to indicate that they preferred this attribute (during the within-subjects part of the survey) [severe thunderstorm scenario: $\chi^2(df = 1, n = 3571) = 2.72, p = 0.10$, Cramer’s $V = 0.03$; flood watch scenario: $\chi^2(df = 1, n = 1863) = 2.07, p = 0.15$, Cramer’s $V = 0.03$; where $df$ is degrees of freedom and $n$ is the sample size]. In other words, repeated exposure to a forecast presentation with the bar attribute did not increase respondents’ tendency to want it. Figure 6 also shows that respondents’ preferences for the end-time text and box attributes (in the severe thunderstorm warning scenario) were much higher than for the bar, with over 90% of respondents indicating they prefer the forecast presentations with each of those attributes than without.

c. Summary of survey 1 findings for informing survey 2: Bar none

The survey 1 results suggest that the bar attribute was not effective overall. Although the bar was helpful for threat identification and was preferred by some, it did not help most respondents accurately identify either the threat start or end time. The end-time text and box each were partially effective at communicating hazardous weather threat information. The end-time text was helpful for threat identification and understanding when a threat expires, and it was highly preferred by respondents. However, providing the end time of a weather threat without corresponding start-time information misled most people into thinking the threats were already in effect, which is problematic for threats that go into effect in the future as in the flood watch scenario. The box, which was evaluated only in the severe thunderstorm scenario, was helpful for threat identification and understanding start time, was attention getting, and was highly preferred by most respondents. However, the box alone did not help respondents accurately identify the end time of the severe thunderstorm threat.

Based on these findings, we refined the set of attributes by omitting the bar, retaining the end-time text but adding start-time text, and retaining the box but also including it in the flood watch scenario. The three attributes were added to each of the weather scenarios and evaluated in survey 2.

4. Results of survey 2

This section presents the survey 2 results in which we tested the start-time text, end-time text, and box
attributes. Although we developed and evaluated experimental forecasts with the end-time and the start-time text attributes separately, the results suggest that these two pieces of information are best coupled to more precisely and fully convey the timing of a weather threat. Therefore, we focus on presenting the results from the experimental forecast presentations that have the combined timing attributes, both with and without the box, alongside the results of the box-only and control forecast presentations. We discuss significant aspects of the results of the other forecast presentations, but full results for every presentation can be found in Demuth et al. (2012a). Results discussed are the same in each weather scenario unless otherwise indicated.

a. Threat identification and timing

As in survey 1, in the between-subjects portion of the survey, respondents were randomly assigned to one of the experimental forecast presentations or to the control. To evaluate threat identification, respondents were asked which type of hazardous weather appeared in the forecast and were offered six response options: the correct weather threat ("severe thunderstorm warning" and "flood watch" for the respective weather scenarios), three incorrect weather threats, "None of the above," and "I don’t know." Ninety percent or more of respondents correctly identified the hazardous weather threat (Fig. 7) illustrating, as in survey 1, that each of the experimental presentations is highly effective at conveying the threat existence when respondents are prompted to examine them. The most effective experimental presentation, however, is the one with the box + start-time text + end-time text. This suggests that these pieces of information interact in ways that maximize drawing respondents’ attention to the threat.

To assess the communication of threat timing, respondents were asked what time the severe thunderstorm warning or flood watch starts and then what time it ends, and they were given five response options: three options with different times,10 "I cannot tell," and "Other." For the severe thunderstorm warning scenario, the correct start and end times were "it has already started (i.e., now)" and "6 PM today." For the flood watch scenario, the correct start and end times were "6 AM on Thursday" and "6 PM on Saturday."

As in survey 1, the control forecasts were the least effective at conveying threat timing (Fig. 8). Respondents who received the control were least likely to accurately identify the start or end time and most likely to respond "I cannot tell" (not shown), again illustrating that the PnC does not convey important threat timing information well.

10Severe thunderstorm warning scenario start-time options were "It has already started (i.e., now)," "6 PM today," and "Sometime tonight"; and end-time options were "It has already ended," "6 PM today," and "Sometime tonight." Flood watch scenario start-time options were "It has already started (i.e., now)," "6 AM on Thursday," and "6 PM on Saturday"; and end-time options were "It has already ended," 6 AM on Thursday," and "6 PM on Saturday."
Although 43% of respondents who received the control correctly identified the severe thunderstorm warning start time, they may be applying their own knowledge that this type of warning is in effect upon issuance.

Just over half of respondents who received the experimental presentation with only the box accurately identified the threat timing (Fig. 8), except for the severe thunderstorm warning end time, for which the modal response (41%) was the more vague response option that the warning would end “sometime tonight” (not shown). These results suggest that many respondents understood that the box represented the timing of the weather threats—especially for the flood watch threat that started and ended after the first forecast period—and they were able to infer a precise time based on the response options offered in the survey. In a real-world situation, however, PnC web page users would not be able to glean the precise threat start or end timing with only the box attribute.

Over 93% of the respondents who received the experimental presentation with only the box accurately identified the threat timing (Fig. 8), except for the severe thunderstorm warning end time, for which the modal response (41%) was the more vague response option that the warning would end “sometime tonight” (not shown). These results suggest that many respondents understood that the box represented the timing of the weather threats—especially for the flood watch threat that started and ended after the first forecast period—and they were able to infer a precise time based on the response options offered in the survey. In a real-world situation, however, PnC web page users would not be able to glean the precise threat start or end timing with only the box attribute.

Over 93% of the respondents who received the experimental presentation with the start-time text + end-time text correctly identified when the hazardous weather threats go into effect and when they expire (Fig. 8). However, the respondents who received the experimental presentation with the box + start-time text + end-time text were even more likely to correctly identify the threat timing (Fig. 8). Additional analysis reveals that 3%–4% more respondents who received the forecast presentation with all three attributes correctly identified both the start and end times compared to those respondents who received the forecast with both timing attributes but not the box.

The above results show that the forecast presentation with the start-time text + end-time text and the forecast presentation with the box + start-time text + end-time text each are highly associated with respondents accurately identifying the threat existence and timing. We further compared these two experimental presentations using data from additional survey questions that asked about respondents’ perceptions of them (Table 2). There were no statistically significant differences between the two experimental presentations in respondents’ perceptions of how well they conveyed the threat start or end times. On the other hand, respondents’ perceptions of the forecast with the box + start-time text + end-time text were statistically significantly more favorable than the forecast without the box at conveying the threat existence and at being more visually appealing and attention getting.

b. Attribute preferences

As in survey 1, next was the within-subjects portion of the survey where respondents were shown a series of a single-attribute experimental forecasts paired with the forecast without that attribute (i.e., the control). For each pair, respondents were asked to indicate which they preferred or whether they had “No opinion/No preference”. As shown in Fig. 9, most respondents indicated they prefer having each of the attributes. Comparing between scenarios, the start-time text was preferred by about 10% more respondents in the flood watch scenario than in the severe thunderstorm warning scenario. Recalling that our survey sample consists of highly experienced users of the PnC web page and, presumably, of NWS hazardous weather products, this may be because some respondents know that severe thunderstorm warnings are in effect when they are issued and thus may not want explicit start-time information added for these

**Fig. 8.** Percentage of survey 2 respondents who identified the correct start and end times by experimental forecast for the (a) severe thunderstorm warning scenario [start time: \( \chi^2(7, n = 3717) = 802.67, p < 0.001, \text{Cramer’s } V = 0.47 \); end time: \( \chi^2(7, n = 3717) = 3047.53, p < 0.001, \text{Cramer’s } V = 0.91 \)] and (b) flood watch scenario [start time: \( \chi^2(7, n = 3747) = 2414.17, p < 0.001, \text{Cramer’s } V = 0.80 \); end time: \( \chi^2(7, n = 3747) = 2521.51, p < 0.001, \text{Cramer’s } V = 0.82 \)].
types of weather threats. However, threats that go into effect at a future time, like flood watches, are those for which the start time is unclear on the NWS PnC web page, which may be why more respondents want the information in these types of scenarios.

5. Summary and discussion

We conducted a multimethod research project aimed at characterizing key strengths and limitations of the NWS point-and-click (PnC) web page from users’ perspectives. From the initial research phase—during which we conducted focus groups, a usability evaluation, and a survey—we found that the existence and timing of hazardous weather threats are not clearly indicated on the PnC web page. Because readily and accurately receiving information about threat existence and timing is critical for PnC web page users to personalize their weather risk, we conducted the study, reported upon here, aimed at more effectively communicating these two aspects. We designed experimental presentations of PnC forecast information by adding new textual and graphical information pieces (i.e., attributes) and evaluated them through two rounds of nationwide surveys of PnC web page users, each with a severe thunderstorm warning and flood watch scenario.

![Table 2](image)

**TABLE 2. Mean responses (on a scale from 1 = not at all to 5 = extremely) to each of the five perception statements shown, for the survey 2 experimental forecast with the start-time + end-time compared with the forecast with the box + start-time + end-time.**

<table>
<thead>
<tr>
<th>Statement</th>
<th>Weather scenario</th>
<th>Mean response: Forecast with start-time + end-time</th>
<th>Mean response: Forecast with box + start-time + end-time</th>
</tr>
</thead>
<tbody>
<tr>
<td>How well does the forecast convey when the (severe thunderstorm warning/flood watch) starts</td>
<td>Severe thunderstorm warning</td>
<td>3.75</td>
<td>3.91</td>
</tr>
<tr>
<td></td>
<td>Flood watch</td>
<td>4.59</td>
<td>4.57</td>
</tr>
<tr>
<td>How well does the forecast convey when the when the (severe thunderstorm warning/flood watch) ends</td>
<td>Severe thunderstorm warning</td>
<td>4.45</td>
<td>4.47</td>
</tr>
<tr>
<td></td>
<td>Flood watch</td>
<td>4.59</td>
<td>4.56</td>
</tr>
<tr>
<td>How well does the forecast convey that a threat of a (severe thunderstorm warning/flood watch) exists</td>
<td>Severe thunderstorm warning</td>
<td>4.34</td>
<td>4.51*</td>
</tr>
<tr>
<td></td>
<td>Flood watch</td>
<td>4.26</td>
<td>4.41*</td>
</tr>
<tr>
<td>This information is visually appealing</td>
<td>Severe thunderstorm warning</td>
<td>3.66</td>
<td>4.04*</td>
</tr>
<tr>
<td></td>
<td>Flood watch</td>
<td>3.72</td>
<td>4.00*</td>
</tr>
<tr>
<td>This information gets my attention</td>
<td>Severe thunderstorm warning</td>
<td>4.05</td>
<td>4.43*</td>
</tr>
<tr>
<td></td>
<td>Flood watch</td>
<td>4.11</td>
<td>4.32*</td>
</tr>
</tbody>
</table>

* Statistically significant ($p \leq 0.01$) differences between mean responses for the two experimental forecasts (within each scenario) based on independent samples $t$-test results.

![Figure 9](image)

**FIG. 9. Percentage of survey 2 respondents who prefer the forecast with the experimental attribute added and who prefer the control forecast (without the added attribute). Results are for the complete samples of respondents who received the (a) severe thunderstorm warning and (b) flood watch scenarios.**
Summarizing the results across the two surveys and two weather scenarios, we found that, first, the control forecast presentation, which had no added attributes and represented the PnC forecast web page as presented at the time of the study, was the least effective overall in helping respondents recognize a threat or understand when a threat was in effect. The control also tended to be perceived less favorably and be less preferred than the other experimental forecast presentations. These results confirm the findings from the initial research phase that improved communication of threat existence and timing on the PnC web page is needed.

Second, we found that the bar attribute, although preferred by many respondents, did not help convey the weather threat start or end time to most respondents, as we had anticipated it would. Had we only elicited respondents' preferences for the attributes and not their understanding of it, the bar may have been deemed useful. These results illustrate the importance of empirically evaluating information and not assuming it will be intuitively understandable to users. It further shows that when deciding how to convey weather risk information, it is important to measure what constitutes effective communication in multiple ways.

Next, we found that the end-time text, as a sole addition to a forecast, was useful for conveying when a threat expires on the PnC web page. However, in both surveys, it significantly increased misunderstandings about the threat start time, leading many respondents to erroneously believe a hazardous weather threat was already in effect. This demonstrates that adding information to address one aspect of the forecast, albeit useful for that purpose, can have unintended negative consequences about another aspect of the forecast. This result again illustrates that it is essential to empirically evaluate people's interpretations of risk information before implementation—particularly when it can significantly impact people's lives and well-being (Demuth et al. 2012b). The American National Standards Institute (ANSI) evaluates risk communication of symbols and warnings in the workplace and for commercial hazards, requiring at least 85% correct interpretation and no more than 5% critical confusion (i.e., where interpretations are opposite of intended) for certification (Lundgren and McMakin 2009). An analogous set of standards for evaluating weather risk communication may be warranted.

When the end-time text was combined with the start-time text in survey 2, together they were useful for helping respondents identify the threat existence and understand the precise threat timing on the PnC web page. In other words, adding the start-time information significantly reduced respondents' confusion about when a hazard product is in effect, which is particularly important for delayed-start threats such as a flood watch. Each of the timing information pieces was also highly preferred by respondents. Because hazardous weather products have specific times that they go into effect and expire, explicitly providing this information as text is a simple, effective way of helping PnC users understand this important risk aspect. Adding the box, which also is highly preferred by most respondents, to the start- and end-time information further enhances communication of a threat. Compared to all other experimental forecast presentations, the combination of these three attributes maximizes drawing attention to the threat existence, precisely conveys the threat start and end times, and is perceived as visually enhancing. In other words, the combined start-time, end-time, and box attributes most effectively communicate weather risk information on the PnC web page based on the measures evaluated here.

Our study focused on two aspects of hazardous weather risk communication on the PnC web page—the existence and timing. Other key aspects of weather risk communication that should be studied include the event location and spatial extent, the event magnitude, the physical and societal impacts of an event, and the uncertainty associated with each of these aspects. Our study also only examined two weather scenarios in depth—one with a rapid-onset, short-duration warning and one with a delayed-start, longer-duration watch. As a first test of other scenarios, we further evaluated respondents' preferences for the box attribute for a severe thunderstorm watch and an urban and small stream advisory, each in the first forecast period. Initial results showed that most respondents preferred having the box on the PnC web page in these scenarios; further, more than half of respondents want the box for all types of hazards (Demuth et al. 2012a).

Another limitation of our study was that the weather scenarios each had only one hazardous weather product in effect during the PnC forecast period. Multiple weather hazards often can threaten a location at the same time or during the multiday forecast period provided on the PnC web page, and even a single type of weather hazard can be associated with multiple products issued for a given location (e.g., a severe thunderstorm watch, a severe thunderstorm warning, a severe weather statement). As a first test of multihazard scenarios, we created a limited number of experimental forecast presentations that included the start-time text, end-time text, and box when multiple hazards existed, including forecasts that had products for single or multiple hazards and forecasts where hazards did and did not overlap in time. Survey respondents' preliminary evaluations of these forecasts generally were positive (Demuth et al. 2012a). Additional research is needed with other single- and multihazard weather scenarios to further examine...
how the start-time text, end-time text, and box attributes are understood, perceived, and preferred in these types of threat situations.

The research conducted here has several potential broader implications. The NWS PnC web page has an average of 2 million unique web hits per day and considerably more during major weather events (B. Akamine, NWS, 2012, personal communication). Looking at Fig. 7, the experimental forecast with the combined start-time text, end-time text, and box increases the threat identification of the severe thunderstorm warning by 4.2% and of the flood watch by 7.8% when compared to the control forecast (which represents the way hazardous weather threats are currently conveyed on the PnC web page). Based on the PnC web page traffic, this could equate to tens of thousands of additional recognitions of hazardous weather threats by PnC users on a given day (depending on the extent of hazardous weather throughout the United States and on the number of people attending, unprompted, to a weather threat in real-world situations; see section 3a). Such additional recognitions can potentially influence if and how these PnC users seek additional information and take protective action. Thus, improvements to the communication of hazardous weather information on the PnC web page—even ones that are small statistically—can translate into very large positive implications practically.

As another example, looking at Fig. 8, the experimental forecast with the start and end times (either with or without the box) increased respondents’ understanding of the threat timing by 50%–95% compared to the control. Without information about when a hazardous weather product goes into effect or when it expires, one may not know that a threat has begun or is still in effect and therefore not know that he or she is at risk. On the other hand, one may consider oneself at risk for a longer period of time than is actually the case; that is, one may be overwarned temporally. Overwarning is a topic of interest and concern in the operational and weather research communities, particularly because it is presumed to lead to warning desensitization (e.g., Mileti and Sorensen 1990; Barnes et al. 2007; Coleman et al. 2011; NOAA 2011) but also because it can have economic implications (Sutter and Erickson 2010). Thus, explicitly indicating when a weather threat is in effect can help people understand and respond during the time period that they are actually at risk.

In summary, the research reported upon here illustrates that, in general, the design of forecast information influences how recipients attend to and interpret the information, which has implications for their weather risk assessment. More specifically, improvements to how weather risk information is conveyed on the PnC web page based on the findings discussed here may substantially enhance decision making and have tremendous positive impacts on people’s lives and well-being. Although our results are directly relevant only to the PnC web page and are generalizable only to the list of PnC users that we recruited, we believe that aspects of the findings are applicable to weather risk communication in other venues.

Acknowledgments. Thank you to Crystal Burghardt, Jennifer Boehnert, Cindy Halley-Gotway, and Taylor Trogdon for their assistance with this research and analysis; to Pat Callahan at ResearchExec for his extensive assistance with the survey design and implementation; and to Greg Guibert, Heather Lazrus, and three anonymous reviewers for helpful comments on the manuscript. Thank you also to the following NWS personnel for their assistance and input throughout this research effort: Bradley Akamine, Bob Bunge, Dennis Cain, Curtis Carey, Sam Contorno, Bob Glahn, Carl Gorski, Andy Horvitz, Eli Jacks, Ron Jones, Suzanne Lenihan, Mark Mitchell, Daniel Nietfeld, and Jennifer Sprague. This work was funded by the NWS Office of Science and Technology and Office of Climate, Water, and Weather Services under Award NA06NWS4670013. Portions of the authors’ time were supported by NCAR’s Collaborative Program on the Societal Impacts and Economic Benefits of Weather Information (SIP), which is funded by the National Science Foundation and the National Oceanic and Atmospheric Administration through the U.S. Weather Research Program under Award NA06OAR4310119. The views and opinions in this paper are those of the authors.

REFERENCES


