Check for updates

DOI: 10.1111/risa.14222

ORIGINAL ARTICLE

The effect of likelihood and impact information on public response to severe weather warnings

Andrea Taylor[1,2] | **Barbara Summers**[1] | **Samuel Domingos**[3,4] | **Natalie Garrett**[5] | **Sophie Yeomans**[5]

[1]Centre for Decision Research, Leeds University Business School, University of Leeds, Leeds, UK

[2]Sustainability Research Institute, School for Earth and Environment, University of Leeds, Leeds, UK

[3]William James Centre for Research, ISPA Instituto Universitário, Lisbon, Portugal

[4]Hei-Lab: Digital Human-Environment Interaction Labs, School of Psychology and Life Sciences, Universidade Lusófona, Lisbon, Portugal

[5]Met Office, Exeter, UK

Correspondence
Andrea Taylor, Centre for Decision Research, Leeds University Business School, University of Leeds, Leeds, UK.
Email: a.l.taylor@leeds.ac.uk

Abstract

Meteorological services are increasingly moving away from issuing weather warnings based on the exceedance of meteorological thresholds (e.g., windspeed), toward risk-based (or "impact-based") approaches. The UK Met Office's National Severe Weather Warning Service has been a pioneer of this approach, issuing yellow, amber, and red warnings based on an integrated evaluation of information about the likelihood of occurrence and potential impact severity. However, although this approach is inherently probabilistic, probabilistic information does not currently accompany public weather warning communications. In this study, we explored whether providing information about the likelihood and impact severity of forecast weather affected subjective judgments of likelihood, severity, concern, trust in forecast, and intention to take protective action. In a mixed-factorial online experiment, 550 UK residents from 2 regions with different weather profiles were randomly assigned to 1 of 3 Warning Format conditions (Color-only, Text, Risk Matrix) and presented with 3 warnings: high-probability/moderate-impact (amber HPMI); low-probability/high-impact (amber); high-probability/high-impact (red). Amongst those presented with information about probability and impact severity, red high-likelihood/high-impact warnings elicited the strongest ratings on all dependent variables, followed by amber HPMI warnings. Amber low-likelihood/high-impact warnings elicited the lowest perceived likelihood, severity, concern, trust, and intention to take protective responses. Taken together, this indicates that UK residents are sensitive to probabilistic information for amber warnings, and that communicating that severe events are unlikely to occur reduces perceived risk, trust in the warning, and behavioral intention, even though potential impacts could be severe. We discuss the practical implications of this for weather warning communication.

KEYWORDS

impact based forecasts, risk communication, risk perception, weather risk, weather warnings

1 | INTRODUCTION

Around the world severe weather poses a threat to lives and wellbeing (Centre for Research on the Epidemiology of Disasters, United Nations International Strategy for Disaster Risk Reduction [CRED, UNISDR] 2018). Effective weather forecast communications that prompt appropriate protective responses are critical to limiting harm from these events (World Meteorological Organization [WMO], 2015), with the need for the development of hazard early warning systems emphasized in the Sendai Framework for Disaster Risk Reduction (United Nations Office for Disaster Risk Reduction [UNDRR], 2015). In order to better align weather warnings with the risk of harm posed by severe weather, meteorological services are increasingly moving away from warnings based on meteorological thresholds alone to

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.
© 2023 The Authors. *Risk Analysis* published by Wiley Periodicals LLC on behalf of Society for *Risk Analysis*.

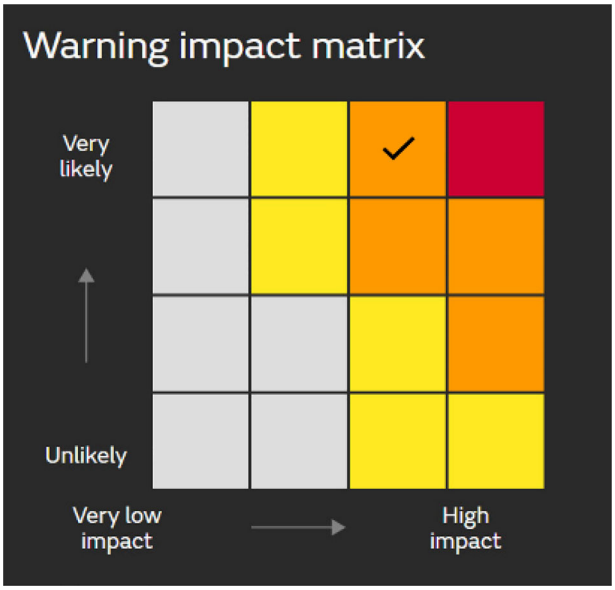

FIGURE 1 Example of the Met Office Risk Matrix for weather warnings.

impact-based warnings, which weight the potential severity of weather impacts by the likelihood of them occurring (WMO, 2015). The UK Met Office has issued impact-based warnings since 2011 (Goldstraw, 2012). However, these are typically provided to the public without explicit information about impact likelihood and severity. This means that the same warning level can be issued for high-probability/low-impact events as for low-probability/high-impact (LPHI) events. Using UK wind warnings, we explore the effect of providing information about likelihood and severity of weather impacts in Text and Risk Matrix formats on perceived risk, forecast trust, and behavioral intention.

1.1 | Severe weather warnings in the UK

In the UK weather warnings are issued by the Met Office's Severe Weather Warning Service for wind, rain, snow, ice, fog, thunderstorms, lightning, and heat (Met Office, 2021a). Yellow, amber, and red denote increasing levels of risk. Warning level is based on a Risk Matrix (Figure 1). Red warnings always indicate high-probability/high-impact (HPHI) events. Yellow and amber warnings, however, can denote varying levels of likelihood and severity. Amber warnings for instance can denote both high-probability/moderate-impact (HPMI) events and LPHI events. Members of the public can access the Risk Matrix on the Met Office website. However, public weather warning dissemination, (e.g., news media, Met Office app), typically presents Color-only warnings. Although the use of the same warning level to capture HPMI and LPHI events follows an expected value approach to risk that weights consequence by likelihood, the characteristics of low, moderate (medium), and high-impact events are different. The behavioral responses required to reduce risk may also differ across impact levels (e.g., being prepared for travel delays for low impact events, cancelling nonurgent journeys for medium-impact events, and sheltering in place for an extended period for a high-impact event) (Met Office, 2021a). If people associate warning levels with impacts alone, they may underestimate or overestimate the potential severity implied by yellow or amber warnings. Moreover, being unaware that a severe event is unlikely but possible may diminish trust in warning providers if it does not materialize (Ripberger et al., 2015). Indeed, LPHI warnings may be issued at longer lead times where there is greater uncertainty about future conditions and subsequently revoked or upgraded to high/medium probability as lead time and uncertainty decrease. Consequently, there are ongoing discussions amongst the operational forecasting community as to how LPHI warnings should be communicated to the public (Zhang et al., 2019).

To date, few studies published in peer-reviewed literature have explored public responses to severe weather warning communications in the UK. Those available suggest that people are broadly aware of the ordinal nature of current color-coded warning levels (Mu et al., 2018; Taylor et al., 2019). However, qualitative work has demonstrated that misinterpretations can occur, particularly with respect to yellow and amber (Tang & Runblad, 2015). In a UK survey exploring public responses to warnings issued for a 2017 severe wind event, Taylor et al. (2019) found that while perceived risk was higher amongst those in areas with amber versus yellow warnings, there was no difference in reported behavioral response. However, noise may have been introduced into this data by some participants incorrectly identifying the warning color for their local area on a warning map. Here, we explore perceived risk and behavioral intention where local warning color is unambiguous, and where warnings are presented with or without probability information.

1.2 | Probability and judgments about weather risk

The utilization of probabilities in decision making has been widely studied in the behavioral sciences. Studies using willingness-to-pay and choice evaluation paradigms have found that sensitivity to probability may be absent or diminished for "affect rich" (i.e., emotionally salient) outcomes such as medical side effects or electric shocks compared to "affect poor" outcomes such as small monetary losses (Rottenstreich & Hsee, 2001; Suter et al., 2016). Finding that willingness-to-pay was less sensitive to probabilities for affect rich (vs. affect poor) descriptions of the same arsenic poisoning event, Sunstein (2003) posited that emotionally salient LPHI risks elicit a focus on impacts at the expense of probability. In the context of weather, where impact severity can vary from minor inconveniences (e.g., traffic delays) to severe outcomes (e.g., fatalities), findings regarding integration of probabilities into decision making have been mixed. In one US study probability neglect was suggested as a

15396924, 2024, 5, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/risa.14222 by University Of Leeds The Brotherton Library, Wiley Online Library on [19/11/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

possible explanation for disparities between perceived risk from flooding and heat and event observations (Allan et al., 2020), although this may have been attributable to recent events biasing responses in line with the availability heuristic (Tversky & Kahneman, 1973). Studies of willingness-to-pay for flood insurance have suggested that while some people are averse to the risk of losses to the extent that likelihoods are neglected, others dismiss low probability events entirely (Botzen & van den Bergh, 2012; Robinson & Botzen, 2019). This may align with prospect theory's weighting function (Kahneman & Tversky, 1979), which holds that small probabilities events may be disregarded (i.e., treated as impossibilities) or overweighted.

In Taylor et al.'s (2019) post-event survey, anticipated impact severity was strongly correlated with anticipated likelihood and trust in warnings. This is consistent with the affect heuristic, whereby feelings toward one characteristic of risk (e.g., potential for harm) may be used as a shortcut to form a generalized assessment of the risk (Finucane et al., 2000). However, as these participants were only provided with the warning color for their region, this may be due to the absence of information on likelihood and severity. Nonetheless, other studies examining perceptions of likelihood and severity have found that when warned for events are more severe they tend to be perceived as more likely to occur (see Ripberger et al., 2022; Weber & Hilton, 1990).

Choice experiments comparing responses to probabilistic versus deterministic weather information have generally been favorable toward the provision of uncertainty information, suggesting that it can lead to better overall decision outcomes over a series of choices (Ramos et al., 2013; Roulston & Kaplan, 2009; Stephens et al., 2019) and reduce loss of trust from false alarms (Joslyn & LeClerc, 2012; LeClerc & Joslyn, 2015). However, providing more information has not universally been found to improve outcomes. Mu et al. (2018) found that while providing risk matrices increased understanding and trust (compared to Color-only warnings), it did not improve scores on a task where participants had to choose whether to take costly protective action or not. However, many of these studies asked participants to take the perspective of an organizational decision maker, rather than consider how they would respond in day-to-day life. Although this provides quantifiable measures of performance, it may not reflect perceptions of personal risk or individual-level behavioral response to severe weather warnings (e.g., changing travel plans).

Here, we explore whether presenting information about likelihood and potential impact severity affects perceived likelihood, potential severity, concern, and behavioral intention. If probability is neglected for LPHI weather events, then providing probability and impact information should lead to greater concern and protective behavioral intention by emphasizing that potential impacts are severe. If communicating that an event is rare but severe leads some to underweight and others to overweight low probability events, then we might expect the standard deviation of response measures to be wider for LPHI than HPMI and HPHI warnings.

1.3 | Trust

In the broader risk perception literature, trust in risk information and information providers has been found to have a complex relationship with risk perception and response, with some studies finding a link between trust and behavioral response (Siegrist, 2021; Wachinger et al., 2013). In the context of weather, trust measures have generally been found to predict greater intention to undertake protective responses (e.g., Kox & Thieken, 2017; Morss et al., 2016; Ripberger et al., 2015; Sherman-Morris, 2005; Taylor et al., 2019). Moreover, there is evidence that providing information about uncertainty may increase trust in warnings and forecasts and reduce a loss of trust from false alarms (Joslyn & LeClerc, 2012; LeClerc & Joslyn, 2015). In the UK and USA it has also been found that warnings for more severe events elicit greater trust than warnings for less severe events (Losee & Joslyn, 2018; Taylor et al., 2019), a finding that may be due to more threatening events creating a psychological need to trust in social systems (Jost et al., 2004; Losee & Joslyn, 2018) but may also be due to trust in forecast reflecting confidence that an event will occur. Indeed, Taylor et al. (2019) found that "trust in forecast" and "trust in the Met Office" as an institution made separate contributions to the prediction of behavioral intention, suggesting that the two are related but distinct.

Here we expand on Taylor et al.'s (2019) work, where only amber and yellow warnings could be compared, to examine trust in red warnings, which represent the highest level of risk (i.e., HPHI). Moreover, we explore whether trust differs among HPHI, HPMI, and LPHI warnings. If severity alone increases trust in forecasts, then trust should be higher for HPHI and LPHI warnings than HPMI warnings. However, if probability affects trust in the forecast, then it should be higher for HPHI than both LPHI and HPMI.

1.4 | Format

This study compares the provision of Color-only warning with those including information about likelihood and impact severity in the form of (a) text description; and (b) the Met Office's Risk Matrix. Multiple studies have investigated the provision of numeric probabilities when communicating about specific meteorological values such as probability of precipitation (e.g., Joslyn & Nichols, 2009; Handmer & Proudley, 2007; Morss et al., 2008). However, in the context of impact-based warnings, attaching numeric values to probabilities and impacts in public-facing communications may not be feasible due to the fact that warnings are not based on exceedance of consistent numeric thresholds (e.g., windspeed exceedance). Presenting information using ordinal categories (e.g., low, medium, high), as captured in the Risk Matrix itself, represents one way to characterize this information non-numerically. Studies examining the use of verbal categories to convey probability and outcome information in hypothetical decision contexts related to health, consumer choice, and environmental management have found

15396924, 2024, 5, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/risa.14222 by University Of Leeds The Brotherton Library, Wiley Online Library on [19/11/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

that they may help people to understand and utilize risk information compared to receiving numeric information alone, through providing an evaluative structure (Dieckmann et al., 2012; Peters et al., 2009). However, a potential disadvantage of using verbal categories is that they may be interpreted differently by different people (Dhami & Mandel, 2021). For instance, interpretation of the terms "very likely" and "likely" has been found to vary considerably between individuals (Budescu et al., 2009). Likewise, terminology related to severity may affect perceptions of probability, with events described as being more severe perceived as more likely (Harris & Corner, 2011).

Risk matrices emphasizing the ordinal nature of "low," "moderate," and "high" categories may remove some ambiguity from interpretation compared to verbal categories and Color-only warnings. Mu et al. (2018) found that risk matrices assisted interpretation relative to Color-only warnings. However, another study examining risk matrices in a broader context did not find that they consistently improved performance on a risk comparison task relative to text controls (Sutherland et al., 2022). In this case, however, both Text and Risk Matrix formats contained numeric likelihood information, which as noted, is unlikely to be feasible in public impact-based warning communications. In this study, we will thus explore whether presenting probability information using verbal descriptions or risk matrices affects perceived risk, trust in forecast, and behavioral intention, in comparison to a Color-only control. If expressing probability and impact verbally elicits varied interpretations, as per Budescu et al.'s (2009) findings, then we would expect to see greater standard deviations in measures of perceived likelihood and severity for the text than the Risk Matrix.

1.5 | Location

In the UK impact-based weather forecasts are regionally calibrated, meaning that the meteorological conditions (e.g., windspeeds) needed to trigger a warning in one area may be different than that for another region, depending on vulnerability and infrastructure (Hemmingway & Robbins, 2020; WMO, 2015). However, many UK residents believe that the conditions needed to trigger a warning are the same everywhere (Taylor et al., 2019). Moreover, there is evidence that perceptions of weather and climate risk differ between southern and northern areas of the UK (Palutikof et al., 2004), with anecdotal evidence from research commissioned by the UK Met Office suggesting that the those in northern areas of the UK may perceive some weather warnings to imply a lower threat to their local area than southern regions (DJS Research, 2014). This raises the concern that the threat implied by a warning may be underestimated when people perceive their region to be resilient to weather events. To empirically test whether this is the case for wind warnings this study purposefully compares responses across two regions: Yorkshire and Humber (Northeast of England), which is heavily exposed to winds from the north and east (Wheeler, 2013); and Greater London (Southeast of England), which has a milder climate, but is vulnerable to the impacts of strong winds due to high population density and infrastructure (Mayes, 2013).

1.6 | Research questions

Based on the areas for investigation identified above, we use an experimental design to address the following research questions:

1. Does perceived likelihood, severity, concern, trust in warning, and behavioral intention differ across warning levels when Color-only wind warnings are presented?
2. Does providing information about likelihood and impact affect the perceived likelihood, impact severity, and concern elicited by wind warnings?
3. Does providing information about likelihood and impact severity affect trust in wind warnings?
4. Does providing information about likelihood and potential impact severity affect behavioral intention for wind warnings?
5. Are there regional differences in wind warning response?

2 | METHODOLOGY

2.1 | Participants

Between August 7 and August 9, 2019 an online experiment was conducted with 550 participants (*female* = 275) from Greater London (n = 275) and Yorkshire and Humber (n = 275). Ages ranged from 18 to 86 (*mean* = 40.2, SD = 14.6). Participants from the two focal regions were recruited from market research panels by Qualtrics Panels (cost = per participant £3.39). Participants received points exchangeable for rewards for their participation. Gender composition was broadly representative of the two regions, though the sample did skew younger than national and regional averages (Office for National Statistics, 2022) (see Supporting Information section for full demographic breakdown).

2.2 | Design

The study used a mixed-factorial design with Location and Warning Format as between-groups factors. Warning level was a repeated measures variable. Participants were randomly assigned to one of three Format conditions (Figure 2):

- Color-only: Statement that a wind warning of a particular color had been issued for participant's local area.

15396924, 2024, 5, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/risa.14222 by University Of Leeds The Brotherton Library, Wiley Online Library on [19/11/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

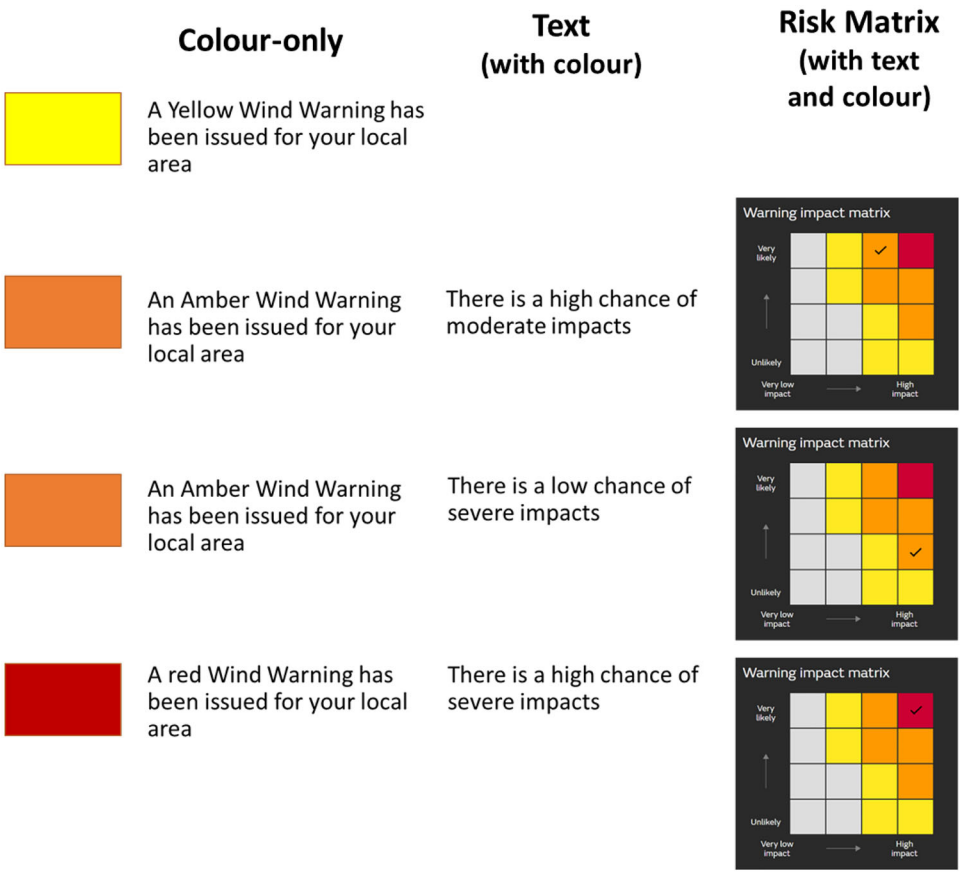

FIGURE 2 Description of each Warning Format condition. Note that the Text condition contains all of the information from the Color-only condition plus descriptive text about likelihood and impact severity, whereas the Risk Matrix condition contains all of the information from the Text condition plus a Risk Matrix visualization.

- Text: Statement that a wind warning of a particular color had been issued for participant's local area, with text description of potential impact severity and likelihood.
- Risk Matrix: Statement that a wind warning of a particular color had been issued for a participant's local area, with text description of potential impact severity and likelihood and a visual representation of the Met Office Risk Matrix.

In the Text and Risk Matrix conditions, participants were shown wind warnings for three warning levels (order randomized): Amber HPMI, amber LPHI event, and red HPMI. As participants in the Color-only condition did not receive any additional information about impact severity or likelihood, they were shown yellow, amber, and red warnings (order randomized).

2.3 | Measures

2.3.1 | Perceived risk

For each warning, slider scales of 0–100 were used to rate expected likelihood of strong wind (0 = impossible, 100 = certain), expected impact severity (0 = not severe at all, 100 = very severe), and expected concern (0 = not concerned at all, 100 = very concerned).

2.3.2 | Trust in forecast

Participants indicated their trust in each warning using a 0–100 slider scale (0 = would not trust at all, 100 = would trust completely).

2.3.3 | Behavioral intention

For each warning, participants indicated how likely they thought that they would be to engage in seven protective behaviors (check for warning updates, notify others, check on vulnerable others, be more cautious when traveling, avoid travel, leave work or study early, take physical action to protect property) on a five-point scale (1 = would definitely not do this, 5 = would definitely do this). Principal components analysis indicated that behaviors loaded onto a single component for all levels (Cronbach's alpha >0.88). A "behavioral intention" score was created by taking the

15396924, 2024, 5, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/risa.14222 by University Of Leeds The Brotherton Library, Wiley Online Library on [19/11/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

mean of all items (see Supporting Information section for information on individual behaviors).

2.3.4 | Trust in the Met Office

Trust in Met Office was measured using the mean of an 11-item scale adapted from Earle and Cvetkovich (1995), presenting pairs of opposing descriptors (e.g., "not trustworthy–trustworthy," "unreliable–reliable") on five-point scales (Cronbach's alpha = 0.91).

2.4 | Analytic approach

Analyses were principally performed using SPSS 27 (IBM), with the PROCESS addon module (Hayes, 2017) for tests of mediation. Pitman–Morgan tests were performed with the PairedData package v.1.1.1 for RStudio (Champely, 2018). Although it would be conventional to use a mixed-factorial analysis to jointly assess the effect of within-groups and between-groups manipulations, it is not appropriate to do this here due to the fact that those in the Color-only condition were presented with yellow, amber, and red warnings, whereas those in the Risk Matrix and Text conditions were presented with amber HPMI, amber LPHI, and red HPHI warnings. Hence, although we present a mixed-factorial ANOVA assessing the interaction between Format and Level for those in the Text and Risk Matrix conditions, separate tests are run for comparisons involving the Color-only condition.

2.4.1 | Warning level

The effect of warning level on perceived likelihood, impact severity, concern, trust in forecast, and behavioral intention is assessed using a repeated-measures ANOVA for the Color-only condition (comparing yellow, amber, and red) and a mixed-factorial ANOVAs for the Text and Risk Matrix conditions (comparing HPMI, LPHI, HPHI across formats). Pitman–Morgan tests are used to compare variability of responses across each level. Using concern as a proxy for perceived personal threat, we assess the extent to which responses of participants in the Color-only condition are aligned with the intended ordinal nature of warnings (i.e., yellow < amber < red), and whether this differs for yellow versus amber than amber versus red, using Wilcoxon sign-rank texts.

2.4.2 | Format and Location

Effect of Format and Location on perceptions of likelihood, impact severity, concern, trust in forecast, and behavioral intention are assessed using MANOVA tests for each warning level. We use Levene's test of equality of error variance to assess whether variability in responses differs across the Text and Risk Matrix conditions (Supporting Information section).

2.4.3 | Predicting behavioral intention

Hierarchical ordinary least squares (OLS) regression is performed with behavioral intention as a dependent variable. Model 1 enters dummy variables representing Location (Yorkshire and Humber as baseline) and Format (Color-only as baseline). Model 2 adds perceived likelihood, severity, concern, trust in forecast, and trust in the Met Office. Mediation tests without covariates are performed to assess whether effects of Location and Format are mediated by concern.

3 | RESULTS

3.1 | Perceived likelihood, severity, concern, trust in warning, and behavioral intention across warning levels for Color-only warnings

Mean ratings of anticipated likelihood, severity, concern, trust in warning, and behavioral intention were highest for red warnings and lowest for yellow warnings, (Table 1). A repeated measures ANOVA with pairwise comparisons demonstrated that across all five dependent variables ratings were significantly higher for red than amber and amber than yellow (Table 1). Standard deviation tended to be higher for yellow warnings than amber and red, suggesting greater variability in responses at the lowest warning level. Pitman–Morgan tests indicated that variance was significantly greater for yellow than amber warnings when it came to perceived likelihood and trust, and greater for yellow than red for trust and behavioral intention.

For the majority of participants the order of reported concern aligned with warning level (Table 2). Few (12%) reported greater concern for yellow warnings than red. However, 26% reported greater concern for yellow than for amber. In comparing whether the frequency with which concern was aligned or not aligned with warning order we found that misalignment was significantly greater for yellow versus amber than yellow versus red ($Z = -4.15$, $p < 0.001$), and marginally significantly greater for yellow versus amber than amber versus red ($Z = -1.93$, $p = 0.053$).

3.2 | Perceived likelihood, severity, concern, trust in warning, and behavioral intention across warning levels for Text and Risk Matrix warnings

Amongst those in the Text and Risk Matrix conditions mean ratings on all dependent variables were higher for HPHI than HPMI warnings, and for HPMI than LPHI warnings (Figure 3, Table 3). These differences were statistically significant (Table 4). As shown in Figure 3, ratings of perceived

15396924, 2024, 5, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/risa.14222 by University Of Leeds The Brotherton Library, Wiley Online Library on [19/11/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

TABLE 1 Ratings of perceived likelihood, severity, concern, trust, and behavioral intention for yellow, amber, and red warnings in the Color-only condition ($n = 184$).

	Yellow warning	Amber warning	Red warning	ANOVA *F*-value	Pairwise-comparisons (Bonferroni-adjusted)	Variances not equal (Pitman–Morgan test)
	Mean (SD)	*Mean (SD)*	*Mean (SD)*			
Perceived likelihood	51.4 (23.1)	57.5 (19.9)	73.1 (20.9)	117.5	Yellow < amber***	A < Y**
					Yellow < red***	
					Amber < red***	
Perceived severity	45.8 (22.9)	54.3 (21.1)	71.3 (22.3)	140.5	Yellow < amber***	–
					Yellow < red***	
					Amber < red***	
Concern	44.2 (25.1)	53.9 (23.4)	70.7 (22.8)	130.1	Yellow < amber***	–
					Yellow < red***	
					Amber < red***	
Trust	57.3 (24.0)	60.4 (21.7)	71.8 (19.3)	60.4	Yellow < amber*	A < Y*
					Yellow < red***	R < Y***
					Amber < red***	R < A*
Behavioral intention	2.8 (1.0)	3.2 (1.0)	3.7 (0.9)	111.3	Yellow < amber***	R < Y*
					Yellow < red***	
					Amber < red***	

Abbreviations: A, amber, R, red; Y, yellow.
*Significant at $p < 0.05$. **Significant at $p < 0.01$. ***Significant at $p < 0.001$.

TABLE 2 Frequency with which ratings of concern aligned or did not align with warning level order of yellow < amber < red.

	Concern aligned with warning order	Concern equal across warnings	Concern not aligned with warning order
Yellow versus amber	126 (68.5%)	10 (5.4%)	48 (26.1%)
Amber versus red	148 (80.4%)	1 (0.5%)	35 (19.5%)
Yellow versus red	155 (84.4%)	7 (3.8%)	22 (12%)

likelihood, severity, concern, and behavioral intention were higher in the Risk Matrix than Text condition for amber warnings (HPMI and LPHI) but not red (HPHI). Consistent with this we found a significant interaction between Format and warning level for each of these (Table 4).

Standard deviations tended to be wider for LPHI warnings than others (Table 3). Pairwise comparisons of equality of variance using Pitman–Morgan tests indicated that differences in variability reached statistical significance ($p < 0.05$) for perceived severity and concern in both conditions, and trust in forecast and behavioral intention in the Text condition. However, as these are repeated pairwise comparisons, the likelihood of a Type 1 error is inflated, meaning that caution is needed in interpreting these findings. For amber warnings standard deviations for the Text condition tended to be wider than those for the Risk Matrix condition (Table 3), Levene's tests indicate that this difference was significant ($p < 0.05$) for trust in forecast (HPMI and LPHI), likelihood (LPHI-only), and concern (HPMI-only) (see Supporting Information section). For red HPHI warnings standard deviations tended to be slightly narrower for the Text than Risk Matrix format, though this only reached significance at $p < 0.05$ for likelihood.

3.3 | Effect of Warning Format and Location on perceived risk, trust in warning, and behavioral intention

Table 5 reports MANOVA tests examining the effect of Location and Format on perceived likelihood, severity, concern, trust in forecast, and behavioral intention for HPMI, LPHI, and HPHI warnings. Responses to the same Color-only amber warning are contrasted with the HPMI and LPHI amber warnings presented for the other conditions. Overall multivariate effects (Wilks Lambda) for Format and Location are reported, along with ANOVAs for each dependent measure, post hoc comparisons for Format and indication of the direction of Location effects (see Supporting Information section for descriptive statistics for Location).

15396924, 2024, 5, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/risa.14222 by University Of Leeds The Brotherton Library, Wiley Online Library on [19/11/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

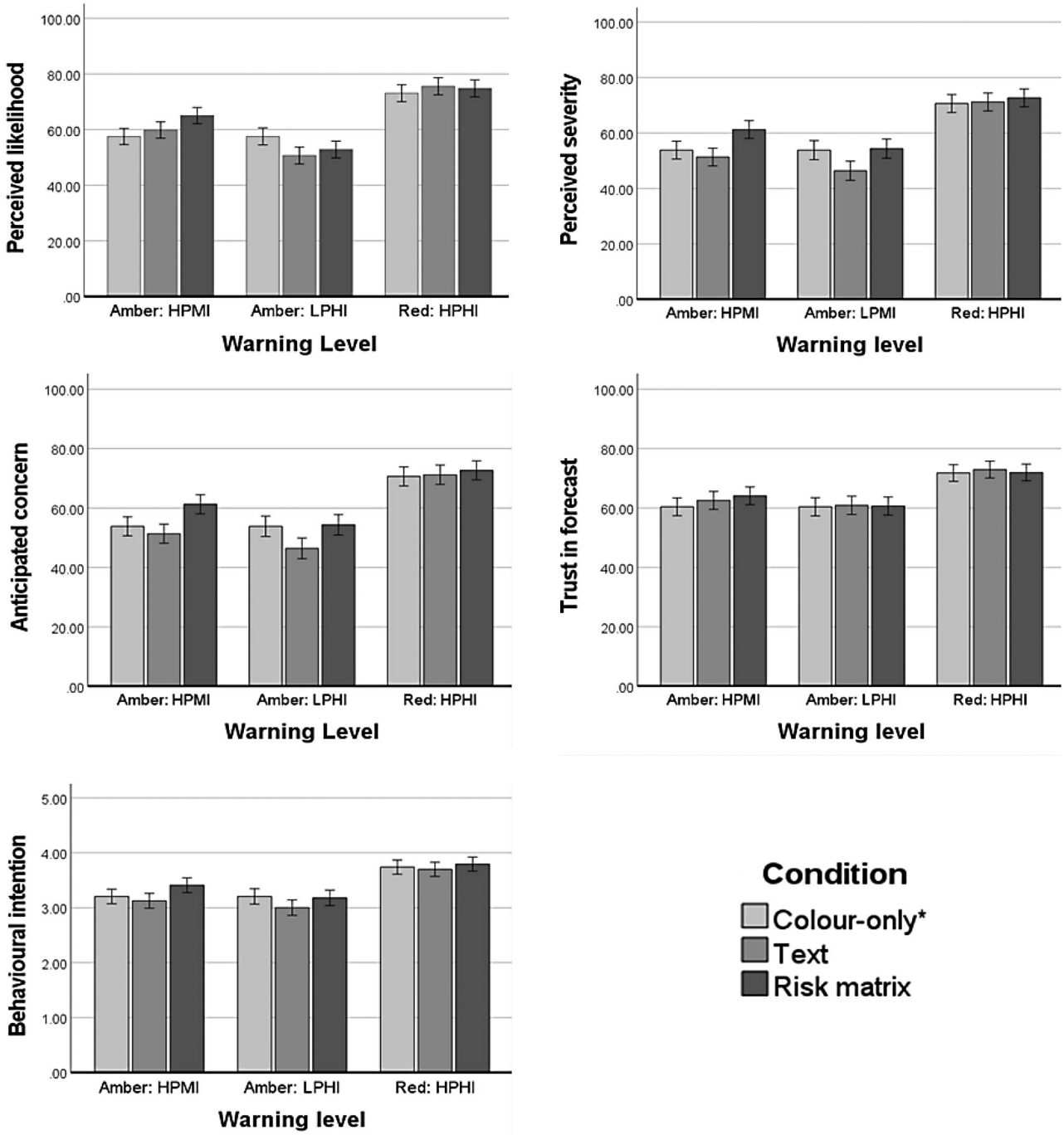

FIGURE 3 Mean ratings of perceived likelihood, severity, concern, trust in forecast, and behavioral intention across each warning level and Format condition. *Note*: Error bars represent 95% confidence interval. *For the Color-only condition responses are identical for amber HPMI and amber LPHI warnings as these participants saw only one amber warning.

No effect of Warning Format was found for responses to red HPHI warnings (Table 5). For the amber HPMI warning the Risk Matrix elicited significantly higher ratings of perceived likelihood, severity, and concern than both the Color-only only and Text formats, and higher behavioral intention than the Text format. For the amber LPHI warning the Risk Matrix elicited greater perceived severity than other formats and greater concern than the Text format, whereas ratings of perceived likelihood and concern were higher for the Color-only than the Text format.

For amber HPMI and LPHI warnings, ratings of concern and behavioral intention were significantly higher amongst those in Greater London than Yorkshire and Humberside (Table 5). No difference in concern and behavior intention was found for red HPHI warnings, though those in Yorkshire and Humberside gave higher ratings of perceived likelihood and severity.

15396924, 2024, 5, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/risa.14222 by University Of Leeds The Brotherton Library, Wiley Online Library on [19/11/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

TABLE 3 Ratings of perceived likelihood, severity, concern, trust, and behavioral intention for warning levels when probability and impact information was provided.

	HPMI (amber)	LPHI (amber)	HPHI (red)	Variances not equal (Pitman–Morgan)
	Mean (SD)	*Mean (SD)*	*Mean (SD)*	
Text: Perceived likelihood	59.9 (20.1)	50.7 (22.6)	75.6 (19.8)	
Risk Matrix: Perceived likelihood	65.1 (20.1)	52.9 (20.1)	74.9 (22.4)	–
Text: Perceived severity	55.7 (20.1)	52.3 (23.2)	74.8 (19.5)	HPMI < LPHI*
				HPHI < LPHI**
Risk Matrix: Perceived severity	63.3 (18.8)	60.0 (21.7)	74.6 (21.2)	HPMI < LPHI*
				HPMI < HPHI*
Text: Concern	51.4 (22.8)	46.4 (25.1)	71.2 (21.6)	HPHI < LPHI*
Risk Matrix: Concern	61.3 (20.0)	54.4 (22.7)	72.7 (22.2)	HPMI < LPHI*
Text: Trust in forecast	62.6 (21.2)	60.9 (22.4)	72.9 (18.7)	HPHI < HPMI*
				HPHI < LPHI**
Risk Matrix: Trust in forecast	64.1 (18.9)	60.7 (18.9)	72.0 (20.1)	–
Text: Behavioral intention	3.1 (0.90)	3.0 (0.98)	3.7 (0.86)	HPHI < LPHI*
Risk Matrix: Behavioral intention	3.4 (0.90)	3.2 (1.00)	3.8 (0.9)	–

Abbreviations: HPHI, high-probability/high-impact; HPMI, high-probability/moderate-impact; LPHI, low-probability/high-impact.
*Significant at $p < 0.05$. **Significant at $p < 0.01$. ***Significant at $p < 0.001$.

TABLE 4 Mixed-factorial ANOVA with pairwise comparisons for ratings of perceived likelihood, severity, concern, trust in forecast, and behavioral intention across warning levels for the Text and Risk Matrix format.

	Warning level (*F*-value)	Format (*F*-value)	Warning level × Format (*F*-value)	Warning level paired-comparisons (Bonferroni corrected)
Perceived likelihood	207.00***	1.60	3.31*	LPHI < HPMI***
				HPMI < HPHI***
				LPHI < HPHI***
Perceived severity	173.18***	8.00**	9.11***	LPHI < HPMI**
				HPMI < HPHI***
				LPHI < HPHI***
Concern	201.98***	10.74***	7.93***	LPHI < HPMI***
				HPMI < HPHI***
				LPHI < HPHI***
Trust in forecast	104.06***	0.004	1.15	LPHI < HPMI**
				HPMI < HPHI***
				LPHI < HPHI***
Behavioral intention	157.84***	4.66*	3.03*	LPHI < HPMI***
				HPMI < HPHI***
				LPHI < HPHI***

Abbreviations: HPHI, high-probability/high-impact; HPMI, high-probability/moderate-impact; LPHI, low-probability/high-impact.
*Significant at $p < 0.05$. **Significant at $p < 0.01$. ***Significant at $p < 0.001$.

3.4 | Predictors of behavioral intention

Stepwise OLS regression analyses examining the predictors of behavioral intention at each warning level are reported in Table 6 (see Supporting Information section for simple bivariate correlations). Model 1, where only Format and Location were entered, was significant for HPMI and LPHI warnings but overall variance accounted for was small (~3.3%). Model 2, which added perceived likelihood, severity, concern, trust in forecast, and trust in the Met Office, accounted for substantially greater variance (>34% for all warning levels). For HPMI and LPHI warnings Location was a significant if small

15396924, 2024, 5, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/risa.14222 by University Of Leeds The Brotherton Library, Wiley Online Library on [19/11/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

TABLE 5 Effect of Location and Warning Format on perceived likelihood, severity, concern, trust, and behavioral intention across warning level (MANOVA with Bonferroni post hoc tests).

	Location		Warning Format	
Amber: HPMI	***Wilks λ = 5.57******		***Wilks λ = 3.70******	
	Main effect *F*	**Direction**	**Main effect *F***	**Post hoc**
Likelihood	0.24(ns)	–	6.27**	RM > CO* RM > T***
Severity	0.86(ns)	–	10.59***	RM > CO*** RM > T***
Concern	11.81***	GL > Y&H	10.17***	RM > CO** RM > T***
Trust	2.45	–	1.46(ns)	–
Behavioral intention	10.56***	GL > Y&H	4.70**	RM > T**
Amber: LPHI	***Wilks λ = 6.37******		***Wilks λ = 5.10******	
	Main effect *F*	**Direction**	**Main effect *F***	**Post hoc**
Likelihood	3.56(ns)		5.06**	CO > T**
Severity	0.14(ns)		6.21**	RM > CO* RM > T**
Concern	14.80***	GL > Y&H	6.40**	CO > T** RM > T**
Trust	1.14(ns)	–	0.03(ns)	–
Behavioral intention	17.18***	GL > Y&H	2.28(ns)	–
Red: HPHI	***Wilks λ = 5.72******		***Wilks λ = 1.00(ns)***	
	Main effect *F*	**Direction**	**Main effect *F***	**Post-hoc**
Likelihood	8.91***	Y&H > GL	0.63(ns)	–
Severity	6.38*	Y&H > GL	1.56(ns)	–
Concern	0.14(ns)	–	0.84(ns)	–
Trust	0.84(ns)		0.18(ns)	–
Behavioral intention	3.18(ns)		0.33(ns)	–

Abbreviations: CO, Color-only; GL, Greater London; HPHI, high-probability/high-impact; HPMI, high-probability/moderate-impact; LPHI, low-probability/high-impact; RM, Risk Matrix; T, Text; Y&H, Yorkshire and Humber.
*Significant at $p < 0.05$. **significant at $p < 0.01$. ***Significant at $p < 0.001$.

predictor of behavioral intention in Model 1, although the association diminished in Model 2. Likewise, the Risk Matrix format was associated with greater behavioral intention for HPMI warnings in Model 1 but not Model 2. In Model 2 concern and trust in the forecast were strongly associated with behavioral intention, with trust in the Met Office making an additional contribution for the model for red HPHI warnings.

The outputs of Models 1 and 2 suggest that the association of Location and Format with behavioral intention was at least partially mediated by variables in Model 2, with earlier MANOVA analyses (Table 5) indicating Location and Format affect concern but not trust in forecast. Formal tests of indirect effects without covariates

(PROCESS Model 4, Figure 4a–c) are consistent with the effects of Location and Format on behavioral intention being mediated by concern. Additional tests controlling for other Model 1 variables yielded the same pattern of findings (see Supporting Information section for full summary of models).

4 | DISCUSSION

4.1 | Does perceived likelihood, severity, concern, trust in warning, and behavioral intention differ across warning levels when Color-only warnings are presented?

Consistent with earlier work suggesting that the ordinal nature of risk implied by color-coded weather warnings is broadly well understood in the UK (Mu et al., 2018; Taylor et al., 2019), our analyses of responses to Color-only warnings indicated that mean ratings of perceived likelihood, severity, concern, and behavioral intention were highest for red warnings and lowest for yellow. However, a notable minority of participants reported higher concern for yellow than amber warnings. In some cases, this may simply be attributable to indifference or error. However, the fact that there were fewer cases where concern was higher for yellow than red does suggest that there may be some ambiguity when it comes to the distinction between the level of

15396924, 2024, 5, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/risa.14222 by University Of Leeds The Brotherton Library, Wiley Online Library on [19/11/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

TABLE 6 Ordinary least squares (OLS) regression examining the predictors of behavioral intention at each warning level.

	Amber HPMI		Amber LPHI		Red HPHI	
Model 1	***β (SE)***	***B***	***β (SE)***	***B***	***β (SE)***	***B***
Location (1 = Greater London)	0.27 (0.08)	0.15***	0.33 (0.08)	0.17***	0.14 (0.08)	0.08
Format: **Color-with-Text	−0.07 (0.10)	−0.04	−0.19 (0.10)	−0.09	−0.04 (0.09)	−0.02
Format: Risk Matrix	0.21 (0.10)	0.11*	−0.03 (0.1)	−0.01	−0.05 (0.09)	0.03
ANOVA *F*	$F(3, 544) = 7.28$		$F(3, 546) = 7.25$		$F(3, 546) = 1.33$	
Adjusted R^2	$R^2 = 0.033$		$R^2 = 0.033$		$R^2 = 0.002$	
Model 2	*β (SE)*	*B*	*β (SE)*	*B*	*β (SE)*	*B*
Location (1 = Greater London)	0.12 (0.07)	0.06	0.16 (0.06)	0.08*	0.11 (0.06)	0.06
Format: Color-with-Text	−0.03 (0.08)	−0.02	−0.04 (0.08)	−0.02	−0.04 (0.08)	−0.02
Format: Risk Matrix	0.06 (0.08)	0.02	−0.05 (0.08)	−0.03	0.03 (0.08)	0.02
Perceived likelihood	−0.004 (0.003)	−0.09	<0.001 (0.002)	0.002	−0.002 (0.003)	−0.04
Perceived severity	0.001 (0.003)	0.02	0.001 (0.002)	0.03	−0.004 (0.003)	−0.08
Concern	0.02 (0.002)	0.49***	0.02 (0.002)	0.52***	0.02 (0.002)	0.48***
Trust in forecast	0.008 (0.002)	0.17***	0.006 (0.002)	0.13**	0.01 (0.003)	0.22***
Trust in Met Office	0.03 (0.05)	0.03	0.01 (0.05)	0.01	0.10 (0.05)	0.08*
ANOVA *F*	$F(8, 538) = 36.45$		$F(8, 541) = 45.97$		$F(8, 541) = 36.60$	
Adjusted R^2	$R^2 = 0.351$		$R^2 = 0.396$		$R^2 = 0.342$	

Abbreviations: HPHI, high-probability/high-impact; HPMI, high-probability/moderate-impact; LPHI, low-probability/high-impact.
*Significant at $p < 0.05$. **Significant at $p < 0.01$. ***Significant at $p < 0.001$.

risk implied by yellow and amber. There was also some indication that responses to yellow warnings tended to be more variable than responses to amber and red warnings, again implying that there may be a lower consensus as to how yellow warnings should be interpreted. A possible explanation for this is that judgments are informed by prior experience of different warning levels. As yellow warnings can indicate a larger number of probability/impact combinations than amber warnings then the experience of the events following them may itself be more variable. In preceding years, participants are likely to have experienced a sizable number of yellow and amber warnings, and infrequent but highly publicized red warnings. This could align with exemplar-based categorization, where a stimulus is compared to stored exemplars of a category (e.g., Erickson & Kruschke, 1998), or the decision by sampling paradigm (Stewart et al., 2006), where experience of events constitutes natural "sampling" of frequency and consequences. However, further research would be needed identify the specific cognitive process(es) through which experience of different colored warning influence subsequent warning risk perception.

The finding that trust in forecast was highest for red warnings and lowest for yellow is consistent with earlier studies indicating that higher warning levels elicit greater trust (Losee & Joslyn, 2018; Taylor et al., 2019). As noted, it has been suggested that this may be due to more threatening situations eliciting a greater need to trust authorities (Jost et al., 2004; Losee & Joslyn, 2018). However, it is also possible that ratings of trust in a forecast reflect participants' confidence that an event will occur given the forecast, rather than beliefs about the quality and credibility of the information per se. Consistent with prior findings that trust in forecasts and trust in forecast providers appear to be related but distinct constructs (Taylor et al., 2019), we find that trust in the Met Office made a unique contribution to the prediction of behavioral intention for red HPHI warnings.

In contrast to earlier work, where no difference in behavioral intention was found between those exposed to yellow and amber wind warnings (Taylor et al., 2019), we found that amber warnings elicited greater intention to engage in protective behaviors than yellow warnings. Differences between these studies may be due to differences in levels of measurement used to record behavioral intention. In Taylor et al. (2019) behavioral intention was coded on a binary scale (Action Taken vs. No Action Taken). Here, participants reported the likelihood of undertaking a range of different actions, potentially capturing greater nuance in intention, and the fact that a protective response may take the form of remaining informed or informing others as well as taking physical actions. Nonetheless, differences in reported behavioral intention for yellow and amber warnings were notably smaller than those between red and amber warnings. This may be attributable to there being a higher number of people who rated yellow warnings as more concerning than amber versus people who rated amber as being more concerning that red. However, it may also reflect greater reported trust associated with red warnings, as well as the comparative rarity and salience of these events (i.e., with red warnings typically being issued no more than twice a year and receiving high media attention). Red warnings may also be interpreted

15396924, 2024, 5, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/risa.14222 by University Of Leeds The Brotherton Library, Wiley Online Library on [19/11/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

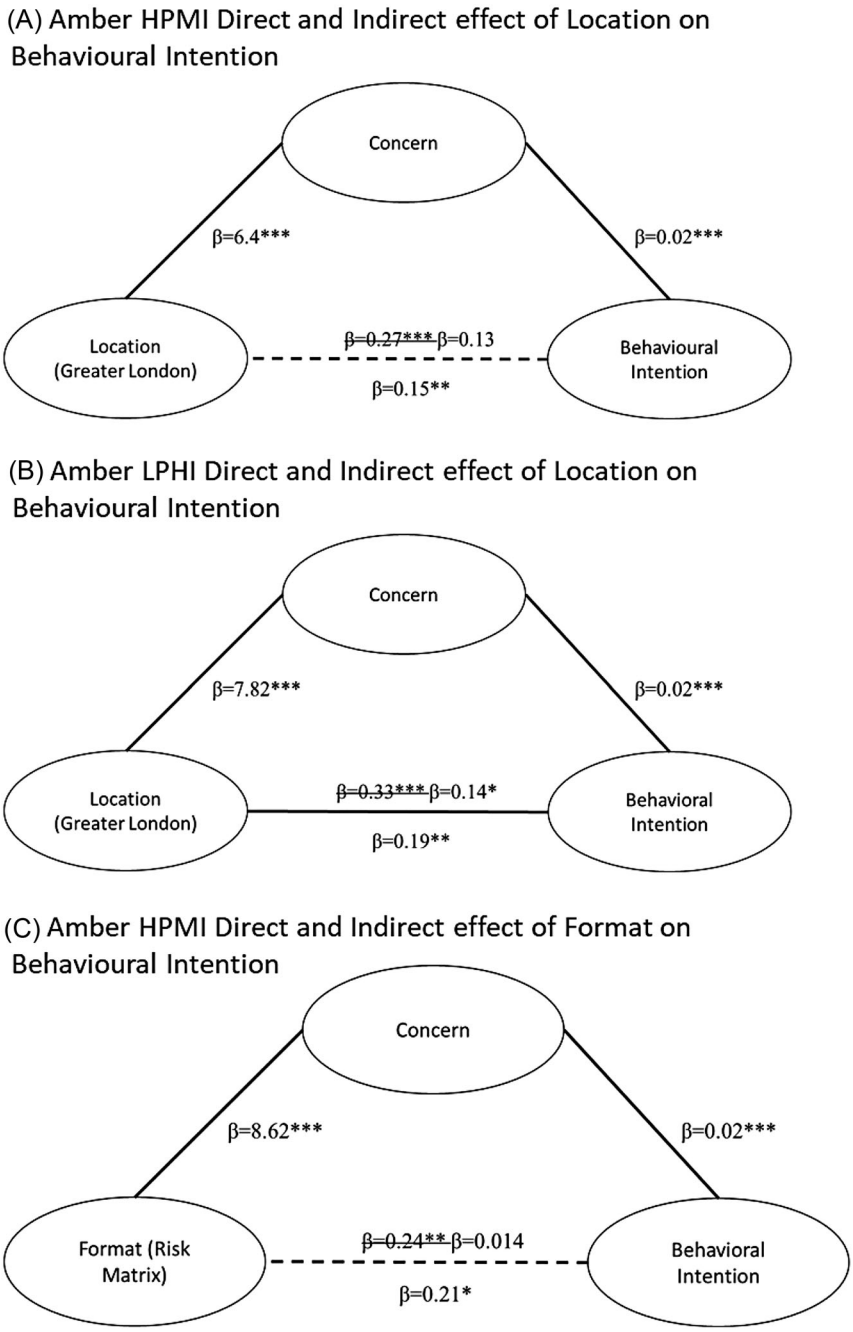

FIGURE 4 (a–c) Tests of direct and indirect effects of Location and Format on behavioral intention. *Note:* $*p < 0.05$, $**p < 0.01$, $***p < 0.001$, β = unstandardized regression coefficient. Total effect of independent variable on behavioral intention is crossed out and reported next to direct effect of independent variable, with indirect effects below these (see Supporting Information section for further statistical details).

as a call to take action in line with phrasing associated with color-coded warnings in earlier UK weather warning messaging (e.g., Neal et al., 2014), which is still used in online Met Office materials describing what a red warning implies (Met Office, 2023).

Taken together these findings demonstrate that without additional information about likelihood and impact severity, red warnings tend to elicit the greatest perceived, trust and willingness to undertake protective behaviors and yellow the lowest. Red warnings, therefore, appear to provide a strong signal that a behavioral response in needed. By contrast, there is some indication that yellow warnings may be perceived as more ambiguous, with higher variability in responses and a notable minority of participants perceiving them to be more concerning than amber warnings.

4.2 | Does providing information about likelihood and impact affect the perceived likelihood, impact severity, and concern elicited by warnings?

In comparing responses across HPMI, LPHI, and HPHI warnings in the Text and Risk Matrix conditions, we find that HPHI warnings consistently elicited higher ratings of perceived likelihood, severity, and concern than both HPMI

15396924, 2024, 5, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/risa.14222 by University Of Leeds The Brotherton Library, Wiley Online Library on [19/11/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

and LPHI warnings. As expected, LPHI warnings elicited lower ratings of likelihood than HPMI warnings. However, they also elicited lower ratings of concern, indicating that the possibility of severe impacts did not lead participants to neglect probabilities. It may be that, for UK residents, the prospect of strong winds associated with amber warnings, which will have been experienced by participants, does not elicit dread in the same way as the as the events addressed by Sunstein (2003). However, the fact that HPHI warnings elicited greater perceived likelihood than HPMI warnings and potential severity than LPHI warnings in the Risk Matrix condition, despite having the same position on the likelihood and impact axis respectively (Figure 2), does suggest that red warnings may have particularly strong affective salience that leads other information to be disregarded or rendered redundant (Leonard, 1999; Silic & Cyr, 2016). This would also align with the salience theory of decision making under risk, which holds that where attention is directed toward one dimension of an outcome (in this case the color red), it may be overweighted relative to less salient outcomes (Bordalo et al., 2012). Indeed, we see that for red (HPHI) warnings responses on all dependent measures were very similar across the Color-only, Text, and Risk Matrix conditions, suggesting that they are always perceived to denote HPHI events irrespective of accompanying information. As the Met Office typically issue no more than two red weather warnings in a year (Met Office, 2021b), and associated events tend to be particularly serious and receive coverage in the news media, this may raise both the cognitive availability and affective salience associated with red warnings.

In considering whether LPHI weather events may be treated in accordance with prospect theory's weighting and editing functions, as suggested by (Botzen & van den Bergh, 2012; Robinson & Botzen, 2019) (i.e., with some underweighting small probabilities and others overweighting them), our analysis does provide some indication that LPHI warnings elicit more variable ratings of concern and severity than other warnings, though no difference in variability was found for perceive likelihood itself. This suggests that there is indeed less consistency in how LPHI warnings are interpreted and responded to. However, it must be noted these differences were relatively small, and whereas the pattern was consistent, the use of repeated pairwise tests did inflate the chance of a Type 1 error.

Although format was not found to affect responses to red warnings, a more complex picture emerges for amber warnings. For perceived likelihood, effects tended to be in the expected directions, with ratings being significantly higher for the Risk Matrix condition than the Color-only condition for HPMI warnings, and significantly lower for the Text versus Color-only condition for LPHI warnings. However, when it came to perceived severity and concern the Risk Matrix elicited the highest ratings across both HPMI and LPHI warnings, whereas Text elicited the lowest. This would seem to indicate that the Risk Matrix increased the affective salience of potential impact severity and associated concern, whereas Text with color attenuated this. As noted, earlier work has highlighted that there may be high variability in how verbal probabilities are interpreted (Budescu et al., 2009; Dhami & Mandel, 2021; Harris & Corner, 2011). Consistent with this, we found some indication that for amber warnings there was greater variability in reported concern amongst those presented with Text versus the Risk Matrix. However, this did not reach statistical significance for severity. It may be the case that the phrasing used in the Text condition was interpreted as ambiguous when not accompanied by the Risk Matrix. That is to say that stating that "there is a low likelihood of severe impacts" may have inadvertently downplayed the possibility of severe impact rather than emphasizing that they may occur Together, this suggests that further work exploring the use of different verbal expressions to capture different magnitudes of probability and impact could be beneficial in examining whether there are better ways to align verbal statements with intended meaning.

4.3 | Does providing information about likelihood and potential severity affect trust in warnings?

As was the case for Color-only warnings, we find that trust in forecast was higher for red warnings than either of the amber warnings for the Text and Risk Matrix. Again, this is consistent with earlier studies indicating that more severe warnings elicit greater trust (e.g., Losee & Joslyn, 2018; Taylor et al., 2019). However, in contrast to earlier work suggesting that providing uncertainty information can increase trust (Joslyn & LeClerc, 2012, 2016) we did not find that providing information about likelihood increased trust in forecast relative to the Color-only condition. Nonetheless, it should be kept in mind that trust in forecast is distinct from trust in forecast providers (Taylor et al., 2019). Here we find that, for HPHI warnings, trust in the Met Office and trust in the forecast provide unique contributions to the prediction of behavioral intention. We therefore postulate that in this case ratings of trust in forecast reflect confidence that an event will occur rather than beliefs about the quality of the information or forecast providers. Indeed, in the broader risk communication literature a conceptual distinction has been drawn between trust and confidence; with the former reflecting perceptions of others' values, and the latter performance (e.g., Siegrist et al., 2003, 2005). Consistent with this, we found that for those provided with likelihood information, trust in forecast was higher for HPMI warnings than LPHI warnings.

Together, our findings show that indicating that weather impacts are likely evokes greater trust in forecasts. Although low "trust in warning" does not necessarily imply low "trust in warning providers," it may contribute to a discounting of warning messages and limit intention to act. As taking protective action in vain can have financial and nonfinancial costs, it is not necessarily incorrect that recipients report low intention to act on low likelihood warnings. However, there may be instances where some form of behavioral response to an unlikely but potentially very severe event would be desirable. As LPHI warnings tend to be issued at longer lead times and then upgraded to HPHI or downgraded to "no

15396924, 2024, 5, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/risa.14222 by University Of Leeds The Brotherton Library, Wiley Online Library on [19/11/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

warning" as lead times decrease, advisable behaviors might involve monitoring weather forecasts for updates or preparing to change plans if needed. In communicating with the public about LPHI events, operational forecasting services should be aware of this, potentially stressing the steps that it is a good idea to take when severe events are possible if unlikely.

4.4 | Does providing information about likelihood and potential severity affect behavioral intention for wind warnings?

In comparing behavioral intention across warning levels, we find that in both the Text and Risk Matrix conditions, behavioral intention was highest for HPHI and lowest for LPHI. The fact that behavioral intention for HPMI warnings was greater than that for LPHI further demonstrates that where probabilistic information is provided for warning warnings it is not neglected, and that this probabilistic information informed behavioral intention. However, when behavioral intention for LPHI and HPMI warnings in Text and Risk Matrix format were compared with Color-only amber warnings, differences did not reach statistical significance. Hence, although behavioral intention may differ when probabilities are explicitly stated to be high versus low, this does necessarily lead to differences in intention from a "no information" scenario. For HPMI warnings, behavioral intention was slightly higher for the Risk Matrix condition than Text condition, with evidence that this is mediated by concern (i.e., with the Risk Matrix eliciting greater concern which in turn lead to higher behavioral intention). Indeed, in keeping with earlier work (Taylor et al., 2019), we found that concern was the strongest predictor of behavioral intention across all warning levels, with trust in forecast also making a consistent contribution.

Based on our findings, we cannot conclusively say that one Warning Format is more effective in prompting protective behavioral intention others. However, where probabilistic information is provided, higher likelihood events correspond with greater behavioral intention. It is therefore important that operational forecasting services be aware that stating that an event is low probability may limit willingness to take precautionary action. As noted, this may not in itself be maladaptive given the costs of taking certain actions in vain. However, it could be useful for forecast communicators to highlight the precautionary behaviors that would be recommended under these conditions.

4.5 | Are there regional differences in warning response?

For red HPHI warnings, no effect of location on concern or behavioral intention was found. For amber HPMI and LPHI warnings, however, concern and behavioral intention were higher for those in Greater London than Yorkshire and Humber, with concern mediating the effect on behavioral intention. Although this is not in itself conclusive proof that those in Yorkshire and Humber perceive strong winds to be less threatening, it is consistent with anecdotal data suggesting that some in northern regions may perceive their area to be more resilient to winter weather impacts than those in the south due to experiencing these events with higher frequency and intensity (DJS Research, 2014). Indeed, this would align with findings in the broader field of risk research indicating that hazards perceived as more familiar can evoke lower perceived risk (e.g., Shavit et al., 2016). As noted, prior work shows that many people in the UK are unaware that warnings take regional vulnerabilities into account (Taylor et al., 2019). Hence, it is possible that warning recipients apply intuitive recalibration to warnings based on perceived vulnerability/resilience of their own location, unaware that this has already been incorporated into the warning. At a practical level, this suggests that more needs to be done to inform public audiences that warnings are regionally calibrated.

4.6 | Limitations and future directions

Although experimental studies have the benefit of permitting randomization and standardization of messaging, a limitation of the approach is that it captures anticipated rather than actual behavior. Tracking behavioral responses to warnings in "real time" could therefore be a logical next step in this area. Indeed, recent Swiss work by Weyrich et al. (2020) has demonstrated the potential of using mobile apps for collecting data on forecast responses. Another possibility would be to longitudinally analyze social media responses to weather warnings across the duration of severe weather events for indicators of engagement with weather informational, as well as cognitive, emotional, and behavioral responses during those events (Domingos et al., 2020; Gaspar et al., 2019, 2021). As seasonality may influence the affective salience of severe weather (Bruine de Bruin et al., 2016; Lefevre et al., 2015), the fact that this study was conducted in the UK summer may have reduced the salience of strong winds, typically associated with winter storms. Longitudinal studies may therefore be useful for capturing the effects of seasonality, as well as enabling comparison of responses to different types of weather event. As noted in Section 4.2, our finding that the Text condition seemed to attenuate concern when it came to LPHI events may have been due to the phrasing chosen. This highlights a critical need for further research on the effect of different types of phrasing to describe probability and impact to ensure that interpretations are aligned with communicators' intentions.

5 | CONCLUSIONS AND IMPLICATIONS

This study has several implications for operational weather forecasting in the UK. Most positively, we find that red warnings are highly trusted, evoking high concern, and behavioral intention, with responses being unaffected by

15396924, 2024, 5, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/risa.14222 by University Of Leeds The Brotherton Library, Wiley Online Library on [19/11/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

region or additional probabilistic information. Hence, in cases where severe impacts are imminent, there is likely to be high public responsiveness to warnings. As noted, this may be attributable to the fact that red warnings can only take on one meaning (high probability of severe impacts), with experience of these events reflecting this. By contrast, amber and yellow warnings—whereas issued more frequently than red warnings—can take on multiple meanings, which are reflected in experience. Indeed, we find responses to yellow warnings, which can take on the greatest number of meanings, tend to have greater variance.

Our findings show that stating severe events are possible but unlikely, diminishes perceived concern, trust in forecast, and intention to act. Care therefore needs to be taken around messaging about such events to avoid risks being discounted, while also recognizing that it may be maladaptive to undertake protective actions carrying high personal cost in response to them. One strategy may be to provide narratives about a "most likely" scenario, while highlighting that more severe conditions are a possibility and explicitly recommend that people should keep abreast of weather updates, as uncertainty regarding whether an event will occur or not tends to diminish as forecast lead times become shorter. Nonetheless, it does highlight the need for more research on how LPHI events can be effectively communicated to avoid both disproportionate worry and dismissal of the risk, especially around the vocabulary that should be used in doing this. Regional differences in concern and behavioral intention between Greater London and Yorkshire and Humber also highlight the need to convey that weather warnings are already regionally calibrated.

In terms of Warning Format, our findings do not allow for clear and unambiguous recommendation about how information about severity and likelihood should be presented. However, for amber warnings we can state that risk matrices appear to increase perceived likelihood, potential severity, and concern, irrespective of where the tick in the matrix is. Whether this heightened perception of risk is desirable or not may be situationally dependent. Ultimately this work provides a step toward better understanding of how people perceive and respond to impact-based weather warnings, and how this may be affected by additional information about likelihood and impact severity. Further exploration of how different approaches to phrasing and visualizations affect response will be important for informing the future development of weather warning communication in the UK.

ACKNOWLEDGMENTS

This work was supported by internal funding from Leeds University Business School's Small Grant Scheme. We thank the World Meteorological Organization's High Impact Weather (HIWeather) project for facilitating this collaboration and providing intellectual support for this study.

ORCID

Andrea Taylor https://orcid.org/0000-0002-8949-1234

REFERENCES

Allan, J. N., Ripberger, J. T., Wehde, W., Krocak, M., Silva, C. L., & Jenkins-Smith, H. C. (2020). Geographic distributions of extreme weather risk perceptions in the United States. *Risk Analysis*, *40*(12), 2498–2508.

Bordalo, P., Gennaioli, N., & Shleifer, A. (2012). Salience theory of choice under risk. *The Quarterly Journal of Economics*, *127*(3), 1243–1285.

Botzen, W. W., & van den Bergh, J. C. (2012). Risk attitudes to low-probability climate change risks: WTP for flood insurance. *Journal of Economic Behavior & Organization*, *82*(1), 151–166.

Bruine de Bruin, W., Lefevre, C. E., Taylor, A. L., Dessai, S., Fischhoff, B., & Kovats, S. (2016). Promoting protection against a threat that evokes positive affect: The case of heat waves in the United Kingdom. *Journal of Experimental Psychology: Applied*, *22*(3), 261.

Budescu, D. V., Broomell, S., & Por, H. H. (2009). Improving communication of uncertainty in the reports of the Intergovernmental Panel on Climate Change. *Psychological Science*, *20*(3), 299–308.

Champely, S. (2018). PairedData: Paired *data analysis*. R package version 1.1.1. https://cran.r-project.org/web/packages/PairedData/index.html

Centre for Research on the Epidemiology of Disasters, United Nations International Strategy for Disaster Risk Reduction [CRED, UNISDR]. (2018). *Economic losses, poverty & disasters, 1998–2017*. Université Catholique De Louvain (UCL). http://dx.doi.org/10.13140/RG.2.2.35610.08643

Dhami, M. K., & Mandel, D. R. (2021). Words or numbers? Communicating probability in intelligence analysis. *American Psychologist*, *76*, 549–560.

Dieckmann, N. F., Peters, E., Gregory, R., & Tusler, M. (2012). Making sense of uncertainty: Advantages and disadvantages of providing an evaluative structure. *Journal of Risk Research*, *15*(7), 717–735.

DJS Research. (2014). National Severe Weather Warning Survey WAVE 8. Internal document of the UK Met Office. Available on request.

Domingos, S., Gaspar, R., Fonseca, H., & Marôco, J. (2020). DeCodeR framework: Data collection and coding for demands and resources appraisal in extreme weather events (DeCodeR framework: Recogida de datos y codificación de exigencias y evaluación de recursos en fenómenos climáticos extremos). *PsyEcology*, *11*(1), 90–103.

Earle, T. C., & Cvetkovich, G. (1995). *Social trust: Toward a cosmopolitan society*. Greenwood Publishing Group.

Erickson, M. A., & Kruschke, J. K. (1998). Rules and exemplars in category learning. *Journal of Experimental Psychology: General*, *127*(2), 107–140.

Finucane, M. L., Alhakami, A., Slovic, P., & Johnson, S. M. (2000). The affect heuristic in judgments of risks and benefits. *Journal of Behavioral Decision Making*, *13*(1), 1–17.

Gaspar, R., Domingos, S., Brito, D., Leiras, G., Filipe, J., Raposo, B., & Telo de Arriaga, M. (2021). Striving for crisis resolution or crisis resilience? The crisis layers and thresholds model and Information and Communication Technology-mediated social sensing for evidence-based crisis management and communication. *Human Behavior and Emerging Technologies*, *3*(1), 40–52.

Gaspar, R., Yan, Z., & Domingos, S. (2019). Extreme natural and man-made events and human adaptive responses mediated by information and communication technologies' use: A systematic literature review. *Technological Forecasting and Social Change*, *145*, 125–135.

Goldstraw, C. (2012). The public weather service customer group. *Weather*, *67*(5), 128–131.

Handmer, J., & Proudley, B. (2007). Communicating uncertainty via probabilities: The case of weather forecasts. *Environmental Hazards*, *7*(2), 79–87.

Harris, A. J., & Corner, A. (2011). Communicating environmental risks: Clarifying the severity effect in interpretations of verbal probability expressions. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *37*(6), 1571.

Hayes, A. F. (2017). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach*. Guilford publications.

Hemingway, R., & Robbins, J. (2020). Developing a hazard-impact model to support impact-based forecasts and warnings: The Vehicle OverTurning (VOT) Model. *Meteorological Applications*, *27*(1), e1819.

15396924, 2024, 5, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/risa.14222 by University Of Leeds The Brotherton Library, Wiley Online Library on [19/11/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Joslyn, S. L., & LeClerc, J. E. (2012). Uncertainty forecasts improve weather-related decisions and attenuate the effects of forecast error. *Journal of experimental psychology: applied*, *18*(1), 126.

Joslyn, S. L., & Nichols, R. M. (2009). Probability or frequency? Expressing forecast uncertainty in public weather forecasts. *Meteorological Applications: A journal of forecasting, practical applications, training techniques and modelling*, *16*(3), 309–314.

Jost, J. T., Banaji, M. R., & Nosek, B. A. (2004). A decade of system justification theory: Accumulated evidence of conscious and unconscious bolstering of the status quo. *Political Psychology*, *25*(6), 881–919.

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, *47*(2), 263–292.

Kox, T., & Thieken, A. H. (2017). To act or not to act? Factors influencing the general public's decision about whether to take protective action against severe weather. *Weather, Climate, and Society*, *9*(2), 299–315.

LeClerc, J., & Joslyn, S. (2015). The cry wolf effect and weather-related decision making. *Risk Analysis*, *35*(3), 385–395.

Lefevre, C. E., Bruine de Bruin, W., Taylor, A. L., Dessai, S., Kovats, S., & Fischhoff, B. (2015). Heat protection behaviors and positive affect about heat during the 2013 heat wave in the United Kingdom. *Social Science & Medicine*, *128*, 282–289.

Leonard, S. D. (1999). Does color of warnings affect risk perception? *International Journal of Industrial Ergonomics*, *23*(5–6), 499–504.

Losee, J. E., & Joslyn, S. (2018). The need to trust: How features of the forecasted weather influence forecast trust. *International Journal of Disaster Risk Reduction*, *30*, 95–104.

Mayes, J. (2013). Regional weather and climates of the British Isles—Part 2: South east England and East Anglia. *Weather*, *68*(3), 59–65.

Met Office. (2021a). *What are the national severe weather warning service impact tables?*. Met Office. https://www.metoffice.gov.uk/weather/guides/severe-weather-advice

Met Office. (2021b). *Past weather events*. Met Office. https://www.metoffice.gov.uk/weather/learn-about/past-uk-weather-events

Met Office. (2023). *Weather warnings guide*. Met Office. https://www.metoffice.gov.uk/weather/guides/warnings

Morss, R. E., Demuth, J. L., & Lazo, J. K. (2008). Communicating uncertainty in weather forecasts: A survey of the US public. *Weather and Forecasting*, *23*(5), 974–991.

Morss, R. E., Mulder, K. J., Lazo, J. K., & Demuth, J. L. (2016). How do people perceive, understand, and anticipate responding to flash flood risks and warnings? Results from a public survey in Boulder, Colorado, USA. *Journal of Hydrology*, *541*, 649–664.

Mu, D., Kaplan, T. R., & Dankers, R. (2018). Decision making with risk-based weather warnings. *International Journal of Disaster Risk Reduction*, *30*, 59–73.

Neal, R. A., Boyle, P., Grahame, N., Mylne, K., & Sharpe, M. (2014). Ensemble based first guess support towards a risk-based severe weather warning service. *Meteorological Applications*, *21*(3), 563–577.

Office for National Statistics. (2022). *Population estimates*. Office for National Statistics. https://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/populationestimates

Palutikof, J. P., Agnew, M. D., & Hoar, M. R. (2004). Public perceptions of unusually warm weather in the UK: Impacts, responses and adaptations. *Climate Research*, *26*(1), 43–59.

Peters, E., Dieckmann, N. F., Västfjäll, D., Mertz, C. K., Slovic, P., & Hibbard, J. H. (2009). Bringing meaning to numbers: The impact of evaluative categories on decisions. *Journal of Experimental Psychology: Applied*, *15*(3), 213.

Ramos, M. H., Van Andel, S. J., & Pappenberger, F. (2013). Do probabilistic forecasts lead to better decisions? *Hydrology and Earth System Sciences*, *17*(6), 2219–2232.

Ripberger, J., Bell, A., Fox, A., Forney, A., Livingston, W., Gaddie, C., Silva, C., & Jenkins-Smith, H. (2022). Communicating probability information in weather forecasts: Findings and recommendations from a living systematic review of the research literature. *Weather, Climate, and Society*, *14*(2), 481–498.

Ripberger, J. T., Silva, C. L., Jenkins-Smith, H. C., Carlson, D. E., James, M., & Herron, K. G. (2015). False alarms and missed events: The impact and origins of perceived inaccuracy in tornado warning systems. *Risk Analysis*, *35*(1), 44–56.

Robinson, P. J., & Botzen, W. W. (2019). Determinants of probability neglect and risk attitudes for disaster risk: An online experimental study of flood insurance demand among homeowners. *Risk Analysis*, *39*(11), 2514–2527.

Rottenstreich, Y., & Hsee, C. K. (2001). Money, kisses, and electric shocks: On the affective psychology of risk. *Psychological Science*, *12*(3), 185–190.

Roulston, M. S., & Kaplan, T. R. (2009). A laboratory-based study of understanding of uncertainty in 5-day site-specific temperature forecasts. *Meteorological Applications: A journal of forecasting, practical applications, training techniques and modelling*, *16*(2), 237–244.

Shavit, T., Lahav, E., & Rosenboim, M. (2016). Don't fear risk, learn about it: How familiarity reduces perceived risk. *Applied Economics Letters*, *23*(15), 1069–1072.

Sherman-Morris, K. (2005). Tornadoes, television and trust—A closer look at the influence of the local weathercaster during severe weather. *Global Environmental Change Part B: Environmental Hazards*, *6*(4), 201–210.

Siegrist, M. (2021). Trust and risk perception: A critical review of the literature. *Risk Analysis*, *41*(3), 480–490.

Siegrist, M., Earle, T. C., & Gutscher, H. (2003). Test of a trust and confidence model in the applied context of electromagnetic field (EMF) risks. *Risk Analysis: An International Journal*, *23*(4), 705–716.

Siegrist, M., Gutscher, H., & Earle, T. C. (2005). Perception of risk: The influence of general trust, and general confidence. *Journal of Risk Research*, *8*(2), 145–156.

Silic, M., & Cyr, D. (2016). Colour arousal effect on users' decision-making processes in the warning message context. In *International Conference on HCI in Business, Government, and Organizations* (pp. 99–109). Springer.

Stephens, E. M., Spiegelhalter, D. J., Mylne, K., & Harrison, M. (2019). The Met Office Weather Game: Investigating how different methods for presenting probabilistic weather forecasts influence decision-making. *Geoscience Communication*, *2*(2), 101–116.

Stewart, N., Chater, N., & Brown, G. D. (2006). Decision by sampling. *Cognitive Psychology*, *53*(1), 1–26.

Sunstein, C. R. (2003). Terrorism and probability neglect. *Journal of Risk & Uncertainty*, *26*(2), 121–136.

Suter, R. S., Pachur, T., & Hertwig, R. (2016). How affect shapes risky choice: Distorted probability weighting versus probability neglect. *Journal of Behavioral Decision Making*, *29*(4), 437–449.

Sutherland, H., Recchia, G., Dryhurst, S., & Freeman, A. L. (2022). How people understand risk matrices, and how matrix design can improve their use: Findings from randomized controlled studies. *Risk Analysis*, *42*(5), 1023–1041.

Tang, C., & Rundblad, G. (2015). The potential impact of directionality, colour perceptions and cultural associations on disaster messages during heatwaves in the UK. *PLoS Currents*, *7*, ecurrents.dis.775c310222d5829cb29b7a414370ca50. https://doi.org/10.1371/currents.dis.775c310222d5829cb29b7a414370ca50. PMID: 25932346; PMCID: PMC4404266.

Taylor, A. L., Kause, A., Summers, B., & Harrowsmith, M. (2019). Preparing for Doris: Exploring public responses to impact-based weather warnings in the United Kingdom. *Weather, Climate, and Society*, *11*(4), 713–729.

Tversky, A., & Kahneman, D. (1973). Availability: A heuristic for judging frequency and probability. *Cognitive Psychology*, *5*(2), 207–232.

United Nations Office for Disaster Risk Reduction (UNDRR). (2015). Sendai Framework for Disaster Risk Reduction 2015–2030. United Nations.

Wachinger, G., Renn, O., Begg, C., & Kuhlicke, C. (2013). The risk perception paradox—Implications for governance and communication of natural hazards. *Risk Analysis*, *33*(6), 1049–1065.

Weber, E. U., & Hilton, D. J. (1990). Contextual effects in the interpretations of probability words: Perceived base rate and severity of events. *Journal of Experimental Psychology: Human Perception and Performance*, *16*(4), 781.

15396924, 2024, 5, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/risa.14222 by University Of Leeds The Brotherton Library, Wiley Online Library on [19/11/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Weyrich, P., Scolobig, A., Walther, F., & Patt, A. (2020). Do intentions indicate actual behaviour? A comparison between scenario-based experiments and real-time observations of warning response. *Journal of Contingencies and Crisis Management*, *28*(3), 240–250.

Wheeler, D. (2013). Regional weather and climates of the British Isles—Part 4: North East England and Yorkshire. *Weather*, *68*(7), 184–190.

World Meteorological Organization (WMO). (2015). *WMO guidelines on multi-hazard impact-based forecast and warning services*. WMO. https://library.wmo.int/doc_num.php?explnum_id=7901

Zhang, Q., Li, L., Ebert, B., Golding, B., Johnston, D., Mills, B., Panchuk, S., Potter, S., Riemer, M., Sun, J., Taylor, A., Jones, S., Ruti, P., & Keller, J. (2019). Increasing the value of weather-related warnings. *Science Bulletin*, *64*(10), 647–649.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Taylor, A., Summers, B., Domingos, S., Garrett, N., & Yeomans, S. (2024). The effect of likelihood and impact information on public response to severe weather warnings. *Risk Analysis*, *44,* 1237–1253. https://doi.org/10.1111/risa.14222

15396924, 2024, 5, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/risa.14222 by University Of Leeds The Brotherton Library, Wiley Online Library on [19/11/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License