

[center title]

Risk seeking in naturalistic decision making from a signal detection perspective

Chao Qin, Susan Joslyn, John Palmer

University of Washington

6/5/26

Please add date to  
title page

JP Comments 6/24/26

In risky situations such as potential severe weather, people make decisions for an uncertain future. For example, when facing a possible tornado, people must decide whether or not to seek shelter before they know if it will be necessary. Taking shelter takes time and, in some cases, resources, but provides needed protection if a tornado occurs. In very dangerous weather situations, people may be advised to take protective action, such as taking shelter, when the probability is as low as <sup>q</sup>10% chance. Under these circumstances, the weather event often fails to occur at the residents' location, referred to as a false alarm, making protective action seem like a waste of time, and potentially increasing reluctance to follow future recommendations. Indeed, research on residents responses to warnings for events such as floods, tornadoes, and hurricanes, suggests that many fail to comply with the recommendations (e.g., taking shelter, evacuation), taking on more risk than necessary, referred to as a risk-seeking tendency (Baker, 1991; McKinley & Urbina, 2008; Parker et al., 2009; Smith & McCarty, 2009; Morss & Hayden, 2010; Nagele & Trainor, 2012; Gibbs & Holloway, 2013; Martin et al., 2017; Rashid et al., 2025).

### **Probabilistic Forecasts**

At present residents do not routinely receive probabilistic information with severe weather warnings, despite the fact that it is encouraged by scholars and increasingly available (National Research Council et al., 2006; Morss et al., 2008; Nagele & Trainor, 2012, Gallo et al., 2016; Karstens et al., 2015). There is experimental evidence that providing probabilistic information helps people better understand the likelihood of the weather event, increases trust in the forecast, and allows them to make better decisions compared to those using a single value or deterministic forecasts (Ash et al., 2014; Joslyn & LeClerc, 2013; Klockow-McClain et al., 2020;

action when the likelihood of severe weather was higher than optimal (e.g., shelter only when the likelihood of a tornado is 50% or greater), referred to as their subjective criterion. Therefore, in the study reported here, we ask whether a higher than optimal subjective criterion, the likelihood above which one decides to take protective action, contributes to the observed risk-seeking tendency. *We refer to this possibility as the biased criteria hypothesis*

A high enough subjective criterion could counteract overestimated subjective likelihood and lead to risk-seeking decisions. For example, suppose the probability of the adverse event is 30% and the economically optimal criterion for choosing the safe option is 20%. A person might choose the risky option in this case even if their perceived likelihood is at, say, 40%, if their subjective criterion is at, say, 50%. Therefore, a biased subjective criterion could explain the combination of overestimated subjective likelihood and risk-seeking decisions. A similar explanation has been used for biases in the motivated reasoning literature. For instance, people require more evidence (higher criterion) to be convinced that their preferred conclusion is false compared to the nonpreferred conclusion (Kunda, 1990; Windschitl et al., 2010).

### Signal Detection Theory and Decision Making

In the work presented here, a signal detection approach was used to ~~isolate~~ <sup>identify</sup> the participants' subjective criterion so that it ~~could~~ <sup>can</sup> be compared to the optimal criterion. Signal detection theory, typically used to understand perceptual processes, concerns two psychological concepts: 1) Separate internal representations of the strength of the stimulus (the to-be-detected signal) and the background in absence of the stimulus (two distributions in Figure 1), as well as 2) the subjective criterion representing the strength of the internal representation of the stimulus

required by the decision maker to identify the it as present (See Figure 1; Macmillan & Creelman, 2005).

Both of the internal representations are assumed to be noisy and follow some distribution. Because of this noise, the same stimulus might be internally represented differently each time it occurs. The ability to differentiate internal representations of the stimulus present versus absent is the person's *sensitivity* in stimulus detection. Sensitivity is determined by the distance between <sup>distributions of the</sup> the two internal representation ~~distributions~~ relative to their variability. The further apart the distributions or the less variability, the greater the sensitivity.

In this context, the subjective criterion is the strength of the internal representation above which the presence of the stimulus is reported (red line in Figure 1). The placement of the subjective criterion indicates whether there is a decision bias towards reporting the presence or the absence of the stimulus. Under the same internal representation, a lower subjective criterion leads to an increase in both hits (reporting stimulus is present when it is) and false alarms (reporting stimulus is present when it is not). In contrast, a higher subjective criterion leads to a decrease in both these outcomes. In short, a change in subjective criterion can lead to a behavioral change even if the internally represented distributions stays the same.

Assuming stochastic internal representations such as those of signal detection theory is not new to decision theory (Busemeyer & Townsend, 1993; Thurstone, 1994; Wallin et al., 2018) and can account for situations in which one does not necessarily make the same decision based on the same external information. This means that, in the context of weather decision tasks, it can explain situations in which one does not take protective action every time the objective probability is above the economically optimal criterion. Furthermore, signal detection theory can be integrated with behavioral economics theories, such as prospect theory that incorporates the

economically optimal criterion. This is a simplifying assumption, as people might change their subjective criterion between decisions depending on the outcome of previous trials (Demnitz & Joslyn, 2020).

*can be in several ways that*  
 This model is used to account for biases in decisions. ~~Biases~~ can be illustrated in Figure 1. *One way is*

a change in the two internal representation distributions on the right of the subjective criterion (hits and false alarms). Larger areas to the right (taking protective action more often) represent a risk-averse bias while smaller areas (taking protective action less) represent a risk-seeking bias.

*also can*  
 Biases ~~could~~ be the result of the subjective criterion, ~~as well as the distribution of subjective likelihood~~. Shifting the subjective criterion to the right of the optimal criterion can lead to a risk-seeking bias while a left-shifted criterion can lead to a risk-averse bias. Inflated subjective likelihood can also cause a right shift (in either one or both internal representations: stimulus present and/or absent) leading to a risk-averse bias. On the other hand, a downward, left shift can lead to a risk-seeking bias.

In addition, a change in sensitivity can also lead to a change in decisions that is not a bias.

A lower sensitivity, where the overlap between the two internal representations increases, ~~lead~~ *This is not a bias also as it affects both distributions.* lead to worse decisions compared to economically optimal. To determine whether our manipulations (see sections below) lead to biases as opposed to this non-bias cause of change in decisions, we also conducted an analysis of sensitivity.

In sum, the study reported here used the random likelihood model to examine the two possible explanations for the tendency toward risk-seeking decisions described above: A bias in the subjective criterion (biased criterion hypothesis) versus a bias in the subjective likelihood (biased likelihood hypothesis).

*These are referred to as the  
 versus the*

## Possible Psychological Mechanisms of Biases in Subjective Criterion

We examined two possible psychological mechanisms that could lead to a shift in the subjective criterion away from the economically optimal criterion. The first is the centering effect, the tendency to bias judgements towards the center of the response range (also called central tendency bias; Poulton, 1979; Olkkonen et al., 2014). Originally observed in perceptual judgements (noise volume, distance, color perception), this effect has been observed in judgement and decision making tasks such as the estimation of opposing bidders' bid in an auction (Poulton, 1979; Radvansky et al., 1995; Olkkonen et al., 2014). It is possible that a similar effect is present in the subjective criterion in decision-making tasks. It can be revealed by manipulating the economically optimal criterion. For instance, an economically optimal criterion at 20% might result in a subjective criterion above 20% closer to the midpoint of 50%, while an economically optimal criterion at 80% might lead to a subjective criterion below 80%, again, closer 50%. It is important to note, however, that the centering effect could also be seen in subjective likelihood.

*Remove  
yellow  
highlight*

A second psychological mechanism that could bias the subjective criterion is the gain-loss framing effect that is central to prospect theory (Tversky & Kahneman, 1979). Risk-seeking in the loss frame might function by raising subjective criterion above optimal while risk-aversion in the gain frame might be due to a lower subjective criterion than is economically optimal. This could be revealed by manipulating the frame (loss/gain). These two psychological mechanisms are not mutually exclusive and might both contribute to a biased subjective criterion.

## Overview of Analyses and Experiments

To begin, two previously published experiments from Qin et al., (2024) were reanalyzed test the random likelihood model and examine the possibility of the framing and centering effects

*and*  
 in subjective likelihood, subjective criterion. Then, two new naturalistic weather decision experiments were conducted to more systematically examine the biased criterion and biased likelihood hypotheses using the random likelihood model.

### **Transparency and Openness**

We report how we determined our sample size, all data exclusions (if any), all manipulations, and all measures in the study, ~~and~~ <sup>the</sup> the study follows JARS (Applebaum et al., 2018). For data availability, hypothesis, and analysis plan of the reanalyses (Qin et al., 2024), please see the transparency and openness sections of the original paper. Analytic codes for reanalyses are available at <https://osf.io/24kxj/>.

For the two new experiments reported below, hypotheses, experimental design, procedure, method, elimination criteria and data analysis plans of were preregistered at the Open Science Framework at <https://doi.org/10.17605/OSF.IO/RUFD8> on Jan. 21st, 2024 (Experiment 1, referred as Experiment 2 in pre-registration) and <https://doi.org/10.17605/OSF.IO/V368N> on March. 29th, 2024 (Experiment 2, referred as Experiment 3 in pre-registration). Registration occurred after data collection but before data analyses. Data and analytic codes are available at the Open Science Framework at <https://osf.io/24kxj/>.

### **Reanalysis of Previous Experiments**

#### **Method**

In the previously published experiments (Qin et al., 2024), participants made decisions, based on tornado warnings, about whether to take shelter (safe option) from possible tornadoes or not (risky option). The tornado warnings were presented either with or without (manipulated between group) a message indicating the probability of a tornado. The reanalysis presented here included only the probabilistic conditions. In these conditions, participants were given reliable

probabilities of the tornado (e.g., 30% chance of tornado) either in the form of a color-coded visualization (red format) or a numeric percentage (tabular format). As the differences between these conditions were not the concern of the analysis reported here, they were combined. There were 83 participants from Experiment 1 and 85 from Experiment 2 in the reanalyses.

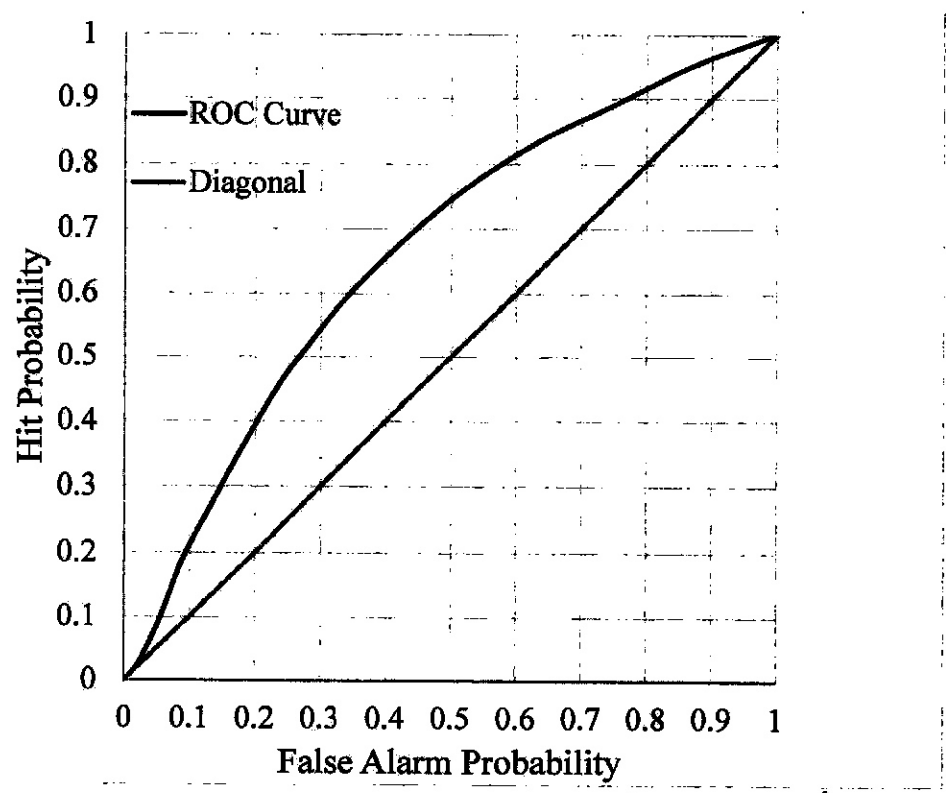
The procedure of the two experiments was identical. There were 68 trials in total. On each trial, participants saw a tornado warning, rated how likely they thought a tornado would be (likelihood ratings), and decided between the safe option and the risky option (binary decision). At the end of each trial, they were told whether the tornado occurred. As mentioned above, there was a point cost to shelter and a larger point penalty if participants failed to shelter and a tornado hit their location. Therefore, only losses were possible. Participants' goal was to lose as few points as possible by the end of the experiment.

The main difference between the two experiments was the ~~point structure~~ (see Table 2). In both experiments, the risky option had no cost but would incur a 1,000-point penalty if a tornado hit. Participants could mitigate this loss completely by choosing the safe option for 90 points in Tornado Experiment 1 and 270 points in Tornado Experiment 2. Therefore, the economically optimal criterion was 9% in Tornado Experiment 1 and 27% in Tornado Experiment 2. Another difference was that in Tornado Experiment 1, 23.5% of trials resulted in a tornado, while in Tornado Experiment 2 the proportion was 38.2%.

## Results

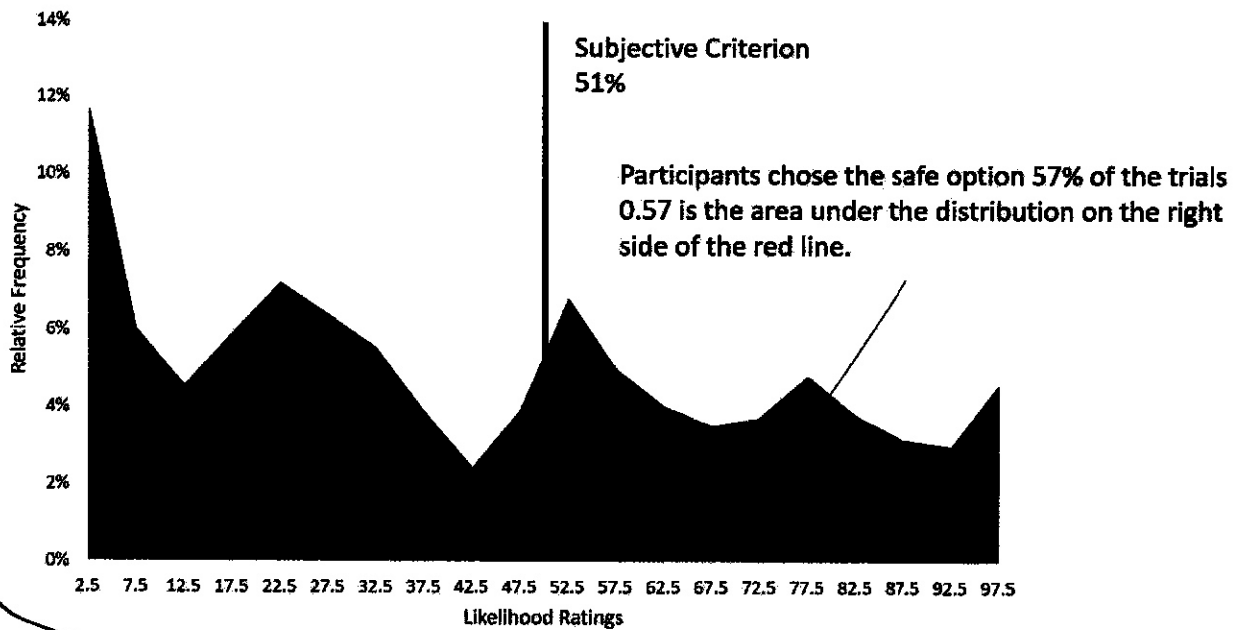
Three sets of analyses were conducted for each experiment. Analyses were conducted on mean subjective criterion and mean likelihood ratings to detect biases. In addition, sensitivity, operationalized as the area under curve in Receiver Operating Characteristic (ROC) plots, were analyzed. ROC plots were constructed from participants likelihood ratings and decisions to

Figure 2 *ROC*  
*A Typical Curve*



Note. The y axis is the probability of a hit *and the x axis is* ~~when the probability of a false alarm is the value on~~  
~~the~~ *axis*. A lower criterion is indicated by an increase in the probability of both a hit and a false alarm, following the blue curve from bottom left to upper right. ✓

Figure 3

*Illustration of a Hypothetical Calculation of Subjective Criterion*

*Note.* The blue area is an example of a likelihood rating distribution.

Make graph text bigger

In Figure 4, the mean subjective criterion for each experiment is shown as a blue dot in relation to the economically optimal criterion. An independent t-test revealed that the mean subjective criterion in Tornado Experiment 1 ( $M = 31.0\%$ ,  $SD = 18.6\%$ ) was significantly lower than in Tornado Experiment 2 ( $M = 41.2\%$ ,  $SD = 13.8\%$ ) with a difference of  $-10.2\%$  (95% CI  $-15.2\%$ ,  $-5.3\%$ ;  $t(148.6) = 4.07$ ,  $p < .001$ , Hedge's  $g = .62$ ) suggesting that the different economically optimal criterion between experiments had an effect on the subjective criterion.

In addition, the mean subjective criterion in each experiment was directly compared to the economically optimal criterion in that experiment using one-sample t-tests. In both cases the mean subjective criterion was higher than the economically optimal criterion. In Tornado Experiment 1, the subjective criterion ( $M = 31.0\%$ ) was significantly higher than the economically optimal criterion of  $9\%$  with a difference of  $22\%$  (95% CI  $18.0\%$ ,  $26.1\%$ ;  $t(82) = 10.77$ ,  $p < .001$ , Cohen's  $d = 1.18$ ). In Tornado Experiment 2, the subjective criterion ( $M = 41.2\%$ ) was also significantly higher than the economically optimal criterion of  $27\%$  with a difference of  $14.2\%$  (95% CI  $11.4\%$ ,  $17.1\%$ ;  $t(84) = 9.81$ ,  $p < .001$ , Cohen's  $d = 1.03$ ). Thus, the subjective criterion was higher than the economically optimal criterion in both experiments, suggesting ~~the~~ <sup>he</sup> ~~risk-seeking biased~~ <sup>was biased to result in risk seeking</sup> subjective criterion. Moreover, the bias was larger in the experiment with the lower economically optimal criterion.

### *Likelihood Ratings*

In each experiment, one-sample t-tests were used to compare the mean likelihood rating over all trials with differing probability levels to the mean objective probability, operationalized as the proportion of trials in which a tornado occurred in that experiment<sup>1</sup>. In both experiments

<sup>1</sup> This is equivalent to the mean objective probability which was reliable.

mean likelihood ratings were higher. In Tornado Experiment 1, the mean likelihood rating ( $M = 33.7\%$ ,  $SD = 10.0\%$ ) was significantly higher than the proportion of tornado trials of 23.5% with a difference of 10.2% (95% CI 8.0%, 12.4%;  $t(82) = 9.33$ ,  $p < .001$ , Cohen's  $d = 1.02$ ). In Tornado Experiment 2, the mean likelihood rating ( $M = 42.7\%$ ,  $SD = 7.5\%$ ) was 4.5% higher than the proportion of tornado trials of 38.2% (95% CI 17.5%, 20.8%;  $t(84) = 5.48$ ,  $p < .001$ , Cohen's  $d = 0.60$ ). Figure 5 shows the likelihood ratings for tornado experiments 1 and 2 as a function of the objective probabilities. It confirms that the <sup>subjective</sup> likelihood was overestimated for all objective probability levels except for the extreme high end. Because the mean objective probability was different between the two experiments (due to different trial composition), it was not possible to compare them directly.

increased with the  
objective likelihood but

### Receiver Operating Characteristic Plots

The ROC analysis was conducted to test the assumption of the random likelihood model that subjective likelihood and the subjective criterion jointly determine the decision. The ~~plots~~ <sup>graphs in Figure 6</sup> were composed of two parts: 1) ROC curves 2) Points indicating the hit and false alarm rates of decisions in each experiment. For a similar approach see Harvey et al. (1992). The calculation of each is explained below. ✓

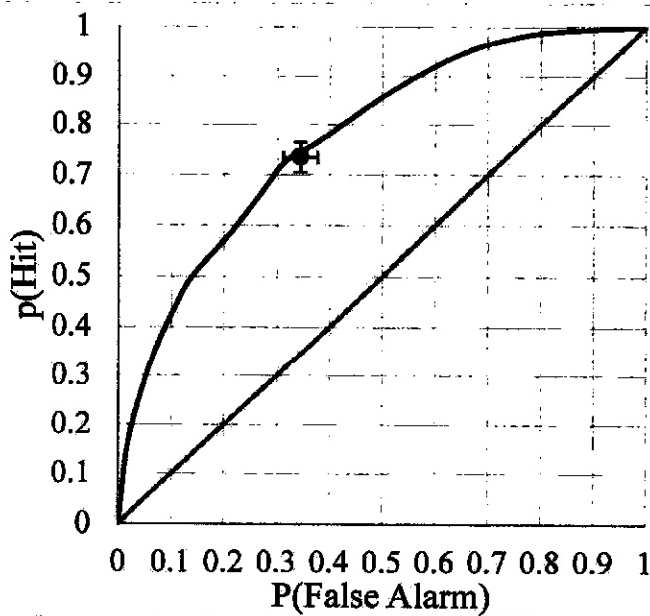
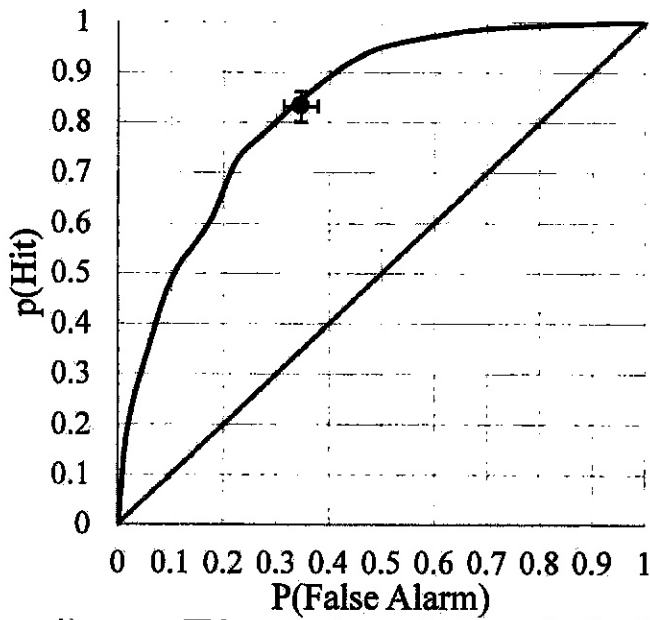
The ROC curves show the relation between the estimated hit and false alarm probabilities at hypothetical cut-off points varying at 5% intervals from 0% to 100% on the likelihood rating distribution (similar to Ferrel & McGoey, 1980). For the curve, a hit was defined as a trial with a likelihood rating above the cut-off ~~and a tornado occurred~~ <sup>with a tornado and</sup>. The hit probability was the proportion of tornado trials above that likelihood cut-off. A false alarm was a trial with a likelihood rating above the hypothetical cut-off ~~and a tornado did not occur~~ <sup>without a tornado and</sup>. The false alarm probability was the proportion of ~~no-tornado~~ <sup>without a tornado</sup> trials above the likelihood cut-off. By varying the cut-off in 5% steps from 0% to 100%, a pair of hit and the false alarm probabilities were calculated at each step (20 total) and plotted as the ROC curve in Figure 6. As 68 trials per participant were too few data points to populate 20 steps to create participant level ROC curves, trials from all participants were combined to create a single ROC curve for each experiment.

Next, a point at the mean proportion of hits (the participant chose the safe option and the tornado occurred) and false alarms (participant chose the safe option and no tornado occurred) in each experiment was added to the appropriate ROC plot.

The first goal of the ROC analysis was to test the internal consistency of the random likelihood model method of calculating the subjective criterion. If participants chose the safe option whenever their likelihood rating was above their subjective criterion, the hit and false

Figure 6

*ROC Plots for Tornado Experiment 1 (Top) and 2 (Bottom)*



Note. The y axis represents the probability of hits. The x axis represents the probability of false alarms. The orange dot indicates the proportion of hits and false alarms for each experiment based on ~~binary decisions~~ <sup>the binary decisions</sup>. Error bars show standard errors of the mean.

## Discussion

In the reanalyses presented above, participants' subjective criterion was higher than the economically optimal criterion in both experiments, indicating a risk-seeking tendency. This analysis does not distinguish whether the bias was due to ~~either~~ framing <sup>or similarity</sup> or centering. Both experiments employed a loss frame which tends to lead to risk-seeking behavior. ~~However,~~ Neither tested a gain frame that would theoretically have produced a relatively risk-averse bias. In addition, the subjective criterion in both experiments was elevated bringing them closer to 50%, consistent with the centering effect. However neither tested a criterion was higher than 50% which, according to centering, would have lowered the subjective criterion creating a risk averse bias. Therefore, it is not possible to distinguish between these two explanations. Hence in the new experiments reported in the next section, the appropriate conditions were added to distinguish between the effects of gain-loss framing and centering.

The reanalysis presented above also suggests that the risk-seeking tendency observed in these two experiments was not due to biased subjective likelihood. The mean likelihood rating was significantly over- rather than underestimated and similar in both experiments. Nevertheless, the higher-than-optimal subjective criterion counteracted the overestimated likelihood ratings and led to risk-seeking decisions<sup>2</sup>.

Finally, in the ROC plots of both experiments, the 95% confidence interval of the binary decision point overlapped with the ROC curve. This indicates that the proportion of hits and false alarms based on participants' binary decisions was consistent with the ROC curve of hits and false alarms based on likelihood ratings. This in turn suggests that the subjective criterion and

---

<sup>2</sup> Due to different trial compositions in the two experiments, <sup>one cannot draw</sup> conclusions about the effect of different economically optimal criterion on likelihood ratings ~~cannot be drawn~~.

loss frame conditions had positive differences between the subjective criterion and the economically optimal criterion (25%) of 6.0% (SD = 14.2%) and 8.8% (SD = 11.7%) respectively. We chose the smaller of the two as our basis, which translates to a Cohen's  $d$  of .43. Based on an alpha of .05, a power of .80, for a two-tailed one-sample t-test<sup>9</sup>, the required sample was 45 for per condition (economically optimal criterion: 25%, 50%, 75%). To account for eliminations due to inattention or incomplete responses, we aimed for a slightly higher number of participants.

### *Participants*

A total of 160 participants from the US were recruited from Prolific Academic in January 2024, a crowdsourcing platform for online research. Each participant received \$4 for participation plus a performance based monetary bonus which ranged from — to —. Three were eliminated for failing the attention check (see Supplementary Materials). 157 participants were used in the analysis.

Demographic data was provided by Prolific. The mean age was 40 (SD = 13.91, range 20 to 80 years). There were 80 (51%) females, 75 (48%) males, 1 (1%) who preferred not to say, and 1 (<1%) with missing data from Prolific. There were 12 Asians (8%), 13 African Americans (8%), 112 Whites (71%), 12 mixed (8%), 6 choosing "other" (4%) and 2 (1%) with missing data from Prolific.

### *Procedure and Stimuli*

The experiment was hosted on Qualtrics. The link to the survey was posted on Prolific inviting those who were residents of the US to participate. The electronic informed-consent form was displayed on the first page and participants were instructed to click next and continue if they

<sup>9</sup>Two-tailed tests were used because the difference for the other two economically optimal criterion conditions might be negative

not necessary

*Post-Task Questions.* After completing all 50 trials, a summary of participants' decisions and outcomes <sup>was</sup> shown along with additional questions. Participants were <sup>first</sup> asked a comprehension check question. Then, as an alternative assessment of criterion, they indicated their (self-reported) decision criterion, in terms of drought probability, for choosing the safe option (drought resistant crop). ~~Participants provided their~~ <sup>THIS</sup> self-reported criterion <sup>was provided</sup> at the end of the experiment <sup>SO IT</sup> ~~which~~ constituted a summary across all 50 trials (see Joslyn & Grounds, 2015).

Because the self-reported criterion was a conscious post-hoc assessment based on memory, it might not match the subjective criterion reflecting both conscious and unconscious trial-specific processes. <sup>if any</sup> Next, participants rated task difficulty and answered an open-ended question asking them which <sup>part</sup> of the experiment they found difficult to understand. These two questions were meant to check the challengingness of the goals in the loss frame task. Finally, participants reported any problems in the experimental program, however, no bugs or glitches were reported. After completion of all questions, participants were thanked and provided with a unique completion code for Prolific to verify their participation and receive payment.

### *Design*

Experiment 1 used a 3 x 6 mixed design. There was one between-group independent variable: Economically optimal criterion with three levels: 25eoc, 50eoc, and 75eoc. There was one within-group independent variable: Objective probability of a drought with five levels: 20%, 35%, 50%, 65% and 80%.

### **Results**

#### *Analysis Overview*

If risk seeking is due to a biased <sup>subjective</sup> criterion and driven by a centering effect, the mean subjective criterion should shift away from the economically optimal criterion towards 50%. This

means that in the 25eoc condition, the subjective criterion ~~would~~ <sup>should</sup> be between 25% and 50% while in the 75eoc condition, the subjective criterion ~~would~~ <sup>should</sup> be between 50% and 75%. In the 50eoc condition, the subjective criterion ~~would~~ <sup>should</sup> be close to 50%. ~~However~~ <sup>In contrast,</sup> the manipulation of the economically optimal criterion should not affect subjective likelihood or sensitivity.

The same set of dependent measures as those used in the reanalyses reported above were analyzed in Experiment 1. These included the subjective criterion and likelihood ratings. As the trial compositions were kept the same across the conditions, an analysis of sensitivity (area under the ROC curve) was conducted as part of the ROC analyses to examine whether participants' ability to predict the drought changed due to the manipulation. For all dependent variables, a series of ANOVAs and t-tests were conducted. Holm-Bonferroni Method was used for planned and post hoc t-tests as well as planned pairwise comparisons under omnibus ANOVAs. Tukey method was used for post hoc pairwise comparisons under omnibus ANOVAs. All tests were two-tailed where applicable.

### *Subjective Criterion*

In Figure 7, the subjective criterion for 25eoc, 50eoc, and 75eoc conditions is shown as blue dots. The mean subjective criterion was 35.0% (SD = 13.3%) in the 25eoc condition, 48.4% (SD = 15.9%) in the 50eoc condition, and 58.1% (SD = 15.3%) in the 75eoc condition. An ANOVA on the subjective criterion with economically optimal criterion (25eoc, 50eoc, and 75eoc) as the independent variable showed a main effect ( $F(2,154) = 32.53, p < .001, \eta_p^2 = .30$ ), suggesting that participants were sensitive to the difference in eoc. In planned pairwise comparisons the 50eoc condition had a significantly higher subjective criterion than the 25eoc condition with a difference of 13.4% (95% CI 7.6%, 19.2%;  $t(154) = 4.54, p < .001$ , corrected alpha = .025, Hedge's  $g = 0.91$ ). The 75eoc condition had a significantly higher subjective

spell out

In addition, the pattern of self-reported criterion (Supplementary Materials S3) was consistent with that of the subjective criterion, although the self-reported criterion was pulled towards 50% to a greater extent than the subjective criterion in the 25eoc condition. <sup>In addition,</sup> The pattern of the proportion of safe option choices was also consistent with <sup>+ the estimated</sup> ~~that of the~~ subjective criterion. <sup>of safe choice</sup> The proportion was lower than economically optimal proportion (risk-seeking) in the 25eoc condition, not different from economically optimal (risk-neutral) in the 50eoc condition, and higher than economically optimal (risk-averse) in the 75eoc condition.

### Likelihood Ratings

Figure 8 shows mean likelihood ratings by eoc condition as a function of objective probabilities. To determine whether the centering effect was observed in perceived likelihood, an ANOVA <sup>was conducted</sup> on mean likelihood ratings collapsed over objective probability levels with the economically optimal criterion manipulation as the independent variable (25eoc, 50eoc, and 75eoc) ~~was conducted~~. It failed to reach significance ( $F(2,154) = 0.81, p = .45, \eta_p^2 = .01$ ). Mean likelihood ratings were similar across conditions (25eoc: 37.1%, SD = 7.0%; 50eoc: 39.3%, SD = 11.9%; 75eoc: 37.8%; SD = 7.1%).

To assess likelihood rating accuracy, post hoc one-sample t-tests compared the mean likelihood rating in each <sup>eoc</sup> condition with the proportion of drought trials (36%). None of the three differences reached significance (25eoc: 1.1%, 95% CI -0.8%, 3.0%;  $t(53) = 1.16, p = .25$ , Cohen's  $d = 0.16$ ); 50eoc: 4.3%, 95% CI -0.1%, 6.7%;  $t(48) = 1.94, p = .06$ , corrected alpha = .017, Cohen's  $d = 0.28$ ); 75eoc: 1.8%, 95% CI -0.1%, 3.8%;  $t(53) = 1.88, p = .07$ , corrected alpha = .025, Cohen's  $d = 0.25$ ). This suggests that the likelihood ratings were fairly accurate and unaffected by the manipulation of economically optimal criterion. In addition, as seen in Figure

and each closely followed the objective probability<sup>40</sup>  
8, the three conditions had similar mean likelihood rating patterns. Likelihood ratings were slightly overestimated at all objective probability levels except for 35%.

### ROC Plots and Sensitivity

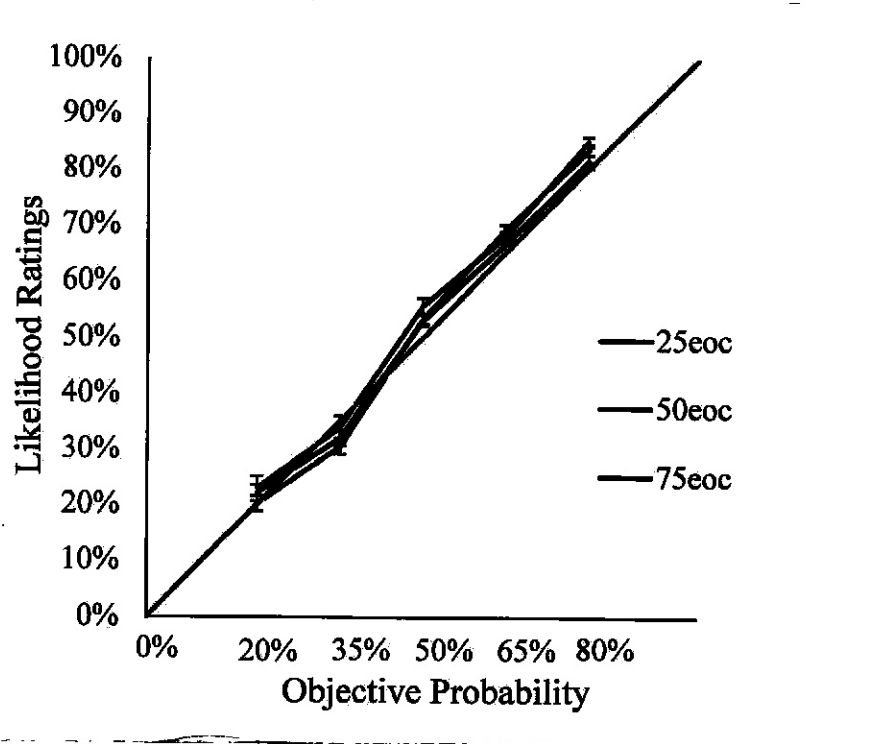
Figure 9 shows ROC plots for the 25eoc, 50eoc, and 75eoc conditions. Notice that the dots representing the mean binary decisions fell on the ROC curve in all three conditions, suggesting that decisions were indeed based on subjective likelihood and subjective criterion, validating the assumption of our analyses.

In addition,  
Also notice that the ROC curves of the three conditions were similar, indicating that all conditions had similar sensitivity (area under the curve). This is a measure of the ability to predict drought based on provided drought forecasts. The mean percent area under ROC curve was 71.2% (SD = 5.0%) in the 25eoc condition, 70.3% (SD = 5.5%) in the 50eoc condition, and 70.6% (SD = 5.4%) in the 75eoc condition. An ANOVA on the mean percent area under curve, with the economically optimal criterion as the independent variable (25eoc, 50eoc, and 75eoc) failed to reach significance ( $F(2,154) = .043$ ,  $p = .65$ ,  $\eta_p^2 = .01$ ). This suggests that the economically optimal criterion manipulation had no observed effect on the sensitivity.

Figure 8

*Mean Likelihood Ratings by eoc condition, as a Function of Objective Probability in Experiment*

1



*Note.* Error bars show standard errors of the mean. Note the similar patterns between lines representing different eoc conditions. Also note that, when comparing the lines representing eoc conditions with the identity line, there is no centering effect.

## Discussion

Experiment 1 tested two predictions based on the on the hypothesis that risk seeking is due to a shift in criterion that shows a centering effect: 1) the subjective criterion <sup>subjective</sup> ~~would~~ <sup>it</sup> shift away from the optimal criterion towards 50%; 2) Likelihood ratings <sup>are</sup> ~~would not be~~ affected by the change in criterion. Supporting the first prediction, the subjective criterion was between 25% and 50% in the 25eoc condition (a shift up towards 50%), and between 50% and 75% in the 75eoc condition (a shift down towards 50%). Meanwhile there was little shift in the 50eoc condition. These results were corroborated by analyses of self-reported criterion and the proportion of trials choosing the safe option in which similar patterns were observed. Thus, in this loss framed task, when compared with the optimal criterion, participants were risk-seeking in the 25eoc condition, risk-neutral in the 50eoc condition, and risk-averse in the 75eoc condition. In addition, as predicted by the biased criterion hypothesis, there was no evidence that manipulating the economically optimal criterion affected the likelihood ratings, ruling biased subjective likelihood out as an explanation for participants biased choices.

Next, the ROC plots (Figure 9) supported the method for calculating participants' subjective criterion. The binary decisions plots were consistent with the ROC curves in each condition, suggesting that the subjective criterion and the subjective likelihood were the sole determinants of the binary decisions. Moreover, sensitivity was unaffected by manipulating the economically optimal criterion.

In summary, Experiment 1 yielded support for the hypothesis that manipulating the economically optimal criterion impacts the subjective criterion but not subjective likelihood or the sensitivity. It is also interesting to note that participants were risk-averse when the economically optimal criterion was high in this loss framed decision task. This hints at the fact

cut?

cut?   
 that the framing effect might not be sufficiently powerful to overcome the centering effect, at least in this scenario. Although it is important to note the framing was not directly manipulated in Experiment 1.

## Experiment 2

The primary goal of Experiment 2 was to manipulate gain-loss framing as well as the economically optimal criterion (centering effect). Experiment 2 used the same drought-related decision task as Experiment 1.

### Method

#### *Power Analysis*

A power analysis, based on a Cohen's  $d$  of .22 (one-tail t-test for the gain- loss frame difference detected in the pilot study), was conducted using G\*Power (Erdfelder et al., 1996). With an alpha of .05 and a power of .80, for a one-tailed independent t-test <sup>for</sup> ~~comparing~~ the difference between the gain and loss frame conditions, the required sample was 268 per condition or 536 for both the gain and loss frame condition. The parameter of one-tailed tests was used as we had a reasonable prediction of the direction of the effect based on the theories and two unreported pilot studies. ✓

#### *Participants*

A total of 591 participants from the US were recruited from Prolific Academic in March 2024. As with Experiment 1, each participant was paid \$4 plus a performance based monetary <sup>that ranged from — to —</sup> bonus. Eleven were eliminated for having a lower than 0.7 ReCAPTCHA score, a bot detection system used by Qualtrics survey platform<sup>5</sup>. Twenty-two were eliminated for failing the comprehension check (same question as Experiment 1, see Supplementary Materials). 558 ✓

<sup>5</sup> Same procedure was used in Experiment 1 with no participant eliminated. ✓

participants were used in the analysis. Demographic data was provided by Prolific. The mean age was 39 (SD = 11.91, range 18 to 81 years). There were 229 (41%) females, 327 (59%) males, 1 (<1%) who preferred not to say, and 1 (1<%) with missing data. There were 62 Asians (11%), 73 African Americans (13%), 359 Whites (64%), 37 mixed race (7%), 24 other (4%) and 3 (1%) with missing data.

### *Procedure and Stimuli*

The procedure, trial structure and stimuli were similar to Experiment 1. Only exceptions are noted here. In particular, a gain frame condition was added (See Table 4) with a drought resistant crop (safe option) yielding a smaller sure gain while the regular crop (risky option) had the potential to yield a higher gain in non-drought conditions (see Supplementary Materials S2). As the centering effect was established in Experiment 1, in order to maximize the number of participants per condition, the less realistic 75eoc was excluded in the second experiment. Severe weather events usually require people to take protective action at a lower probability. In terms of dependent variables, the same questions were asked as with Experiment 1 (See Supplementary Materials S2).

*Point Structure.* The point structure in the 25eoc and 50eoc loss frame conditions was identical to that of Experiment 1 while the point structure in the gain frame conditions had the equivalent risk (See table 4). The beginning point balance and payment structure were the same as the 25eoc and 50eoc conditions of Experiment 1.

## *Design*

Experiment 1 used a 2 x 2 x 6 mixed design. There were two between-groups independent variables: Economically optimal criterion with two levels: 25eoc, and 50eoc and gain-loss framing with two levels: Gain and loss frame. There was one within-group independent variable: objective probability of a drought with five levels: 20%, 35%, 50%, 65% and 80%.

## **Results**

### *Analysis Overview*

As with Experiment 1, we expected the results to indicate a biased subjective criterion due to gain-loss framing and/or centering effects. The effect of framing is supported if there is a higher subjective criterion in the loss than in the gain frame condition, signaling a greater risk-seeking tendency. As with Experiment 1, the centering effect should shift the subjective criterion in the 25eoc condition up towards 50% while having no effect on the subjective criterion in the 50eoc condition. No interaction between these two effects was expected as their mechanisms are theoretically independent. Neither manipulation was expected to affect likelihood ratings or sensitivity. All tests were two-tailed where applicable. ✓

### *Subjective Criterion*

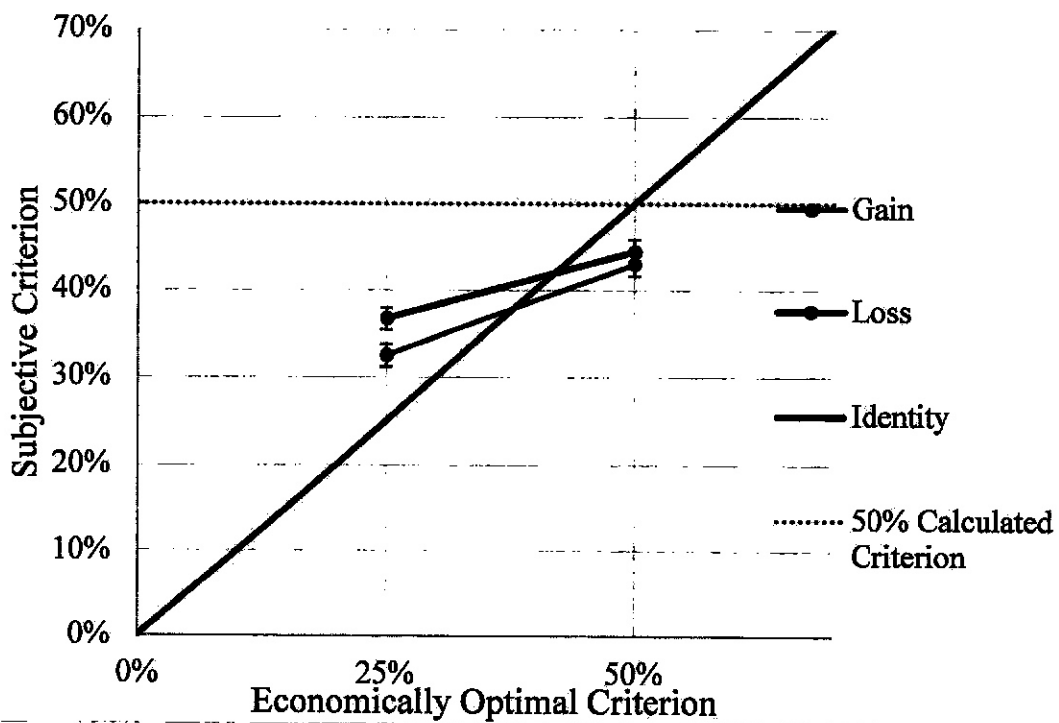
Figure 10 shows the subjective criterion for the four conditions. An ANOVA on the subjective criterion with the gain-loss framing (gain frame, loss frame) and economically optimal criterion (25eoc, 50eoc) revealed a main effect of the gain-loss framing manipulation such that in the loss frame ( $M = 40.6\%$ ,  $SD = 15.9\%$ ) the subjective criterion was 2.7% higher than in the gain frame ( $M = 37.9\%$ ,  $SD = 16.9\%$ ) as predicted by prospect theory ( $F(1, 554) = 4.38$ ,  $p = .037$ ,  $\eta_p^2 = .01$ ). There was also a main effect of manipulating the economically optimal

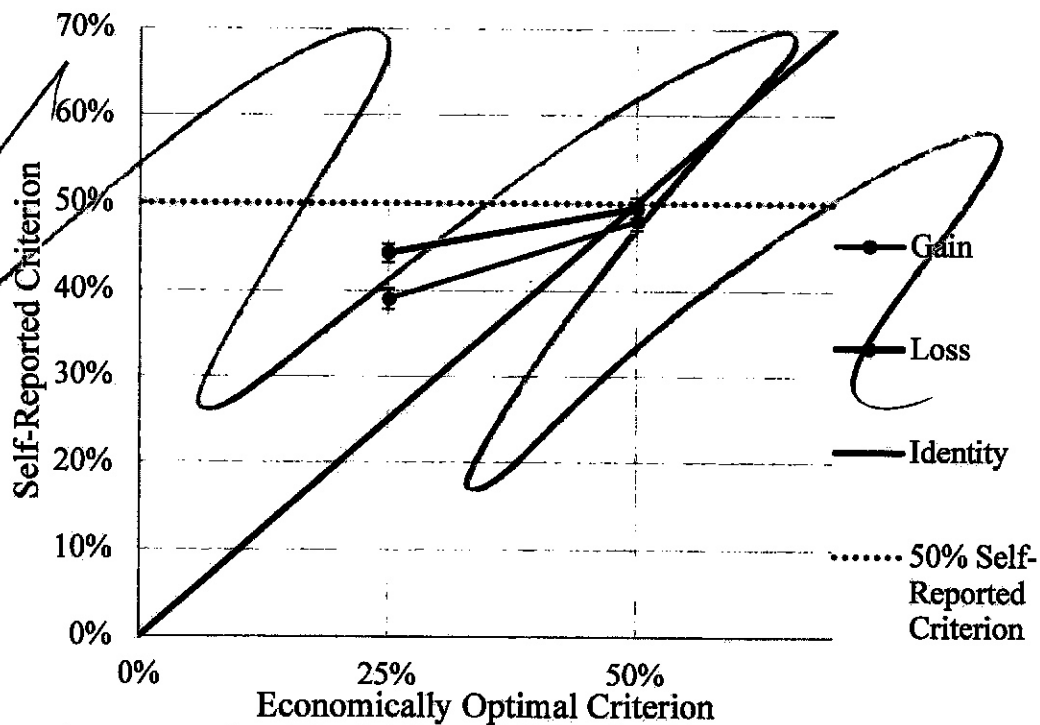
In summary, both manipulations affected the subjective criterion and they did not interact <sup>significantly</sup> with one another. While the subjective criterion in the 25eoc condition was higher than the economically optimal criterion as predicted, the subjective criterion in the 50eoc condition was unexpectedly lower than the economically optimal criterion. ✓

In addition, like Experiment 1, the pattern of self-reported criterion (See Figure 10) and proportion choosing the safe option was consistent with that of the subjective criterion <sup>pattern of results with the</sup> (Supplementary Material S4). ^

Figure 10

*Subjective (top graph) and Self-Reported (bottom graph) Criterion in the Four Conditions of Experiment 2*

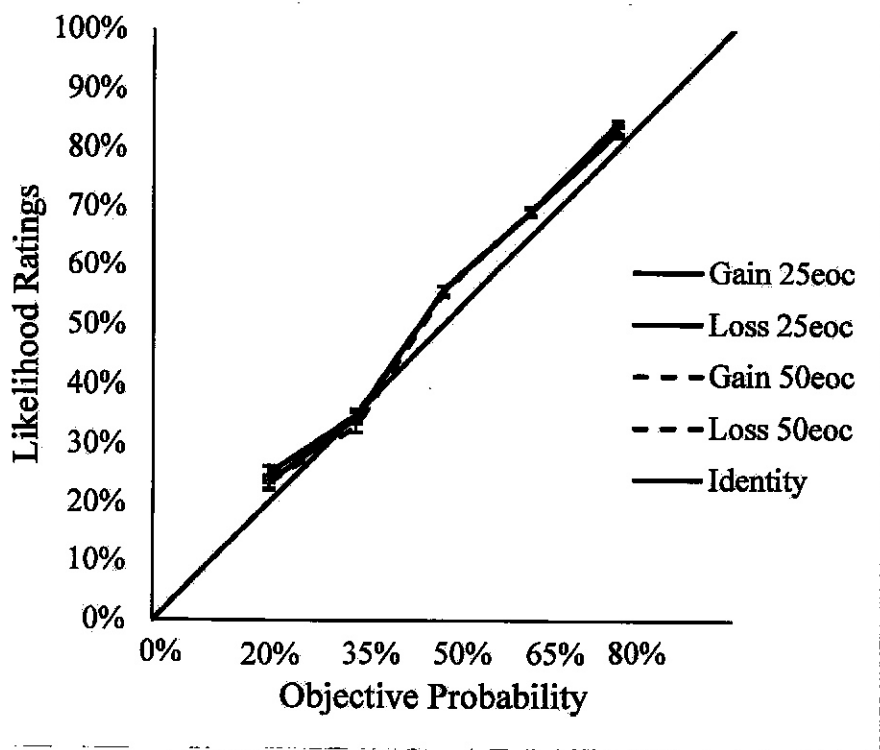




*Note.* The y axis represents the subjective criterion while the x axis represents the economically optimal criterion. The grey diagonal line represents when the subjective criterion is the same as the economically optimal criterion. The grey dashed line represents a subjective criterion of 50%. Error bars show standard errors of the mean.

Remove 2<sup>nd</sup> graph

Figure 11

*Likelihood Ratings a Function of Objective Probability in Experiment 2*

Note. Error bars show standard errors of the mean. Note the similar patterns between lines representing different eoc conditions. Also note that, when comparing the lines representing eoc conditions with the identity line, there is no centering effect.

The likelihood rating increased with the objective probability but showed no effect of the different EOC and gain-loss conditions

## ^ Sensitivity

The next analysis examined sensitivity (participants' ability to predict drought), as measured by the area under the ROC curve. See Figure 12. In an ANOVA on the mean percent area under the curve with the gain-loss framing (gain frame, loss frame) and economically optimal criterion (25eoc, 50eoc) as the independent variables, neither the main effect of gain-loss framing ( $F(1, 554) = 0.51, p = .47, \eta_p^2 < .001$ ) nor that of economically optimal criterion reached significance ( $F(1, 554) = 1.93, p = .17, \eta_p^2 < .01$ ). The interaction also failed to reach significance ( $F(1, 554) = 0.94, p = .33, \eta_p^2 < .01$ ). This suggests that sensitivity was not affected by the manipulation of gain-loss framing or economically optimal criterion. The mean percent Area under ROC curve was 70.0% (SD = 8.1%) in the 25eoc gain, 71.1% (SD = 5.2%) in the 25eoc loss, 69.8% (SD = 7.6%) in the 50eoc gain, and 69.6% (SD = 6.9%) in the 50eoc loss condition. In addition, the 95% confidence interval of binary decisions overlapped with the ROC curves in all but the 50eoc gain frame condition suggesting that participants' decisions can be explained by the combination of subjective likelihood and subjective criterion.

## Discussion

Experiment 2 tested the hypotheses that two ~~mechanisms~~ <sup>manipulations</sup> lead to a biased criterion but not biased subjective likelihood: 1) The gain-loss framing effect, in which the subjective criterion tends to be higher in the loss than in the gain frame. 2) The centering effect, in which the subjective criterion migrates toward the center of the range (50%). Both <sup>effects</sup> were observed and did not significantly interact with one another. Nor did they impact likelihood ratings. } reverse order

The mean subjective criterion was higher than 25% in the 25eoc condition suggesting movement toward the center (50%). The difference between the subjective criterion and the economically optimal criterion was greater in the 25eoc condition than in the 50eoc condition where little movement was expected. Interestingly, the subjective criterion in the 50eoc condition was slightly but significantly lower than 50%. This is corroborated by the proportion of safe choices in the 50eoc condition, where participants tended to be risk-averse (Supplementary Materials S3). This result was also consistent with the ~~inverse~~ trend in proportion of safe choices observed in Experiment 1. This could mean either that the center in the centering effect was slightly lower than 50%, or some other factors shifted the subjective criterion down.

There was also evidence of a framing effect. The subjective criterion in the loss frame condition was significantly higher than that in the gain frame condition. This resulted in participants choosing the safe option less often and showing a risk-seeking decision bias in the loss compared to the gain frame.

It is also interesting ~~that~~ that the effect sizes for the two manipulations differed in the 25eoc condition which is most similar to real world situations for which precautions are warranted for low-probability high-impact weather events. Gain-loss framing shifted the subjective criterion by 4.2% in the 25eoc condition while the centering effect shifted the

subjective criterion to 9.5% above 25%. Therefore, at the 25% economically optimal criterion, the centering effect had a larger impact on the subjective criterion than the gain-loss framing effect, which may have implications for applied communication interventions.

Consistent with Experiment 1, the likelihood ratings were not affected by manipulating economically optimal criterion. This experiment also found that the ratings were not affected by gain-loss framing. However, they were overestimated compared to the objective probabilities as shown in Figure 11. This is also consistent with Experiment 1 and prior research that suggests that biased subjective likelihood did not contribute to the risk-seeking behavior observed in the loss frame.

~~MMI~~ In the ROC plots, the binary decisions (summarized by proportion of hits and false alarms in decisions) were consistent with the respective ROC curves (from likelihood ratings conditionalized by drought occurrence) in almost all conditions suggesting that decisions were jointly determined by subjective likelihood and subjective criterion. ~~MMI~~ In the 50eoc gain frame condition alone, which was not tested in previous experiments, the binary decisions were below the ROC curve. This is the only inconsistency in the ROC analysis observed across both experiments. It is unknown whether this inconsistency was a statistical false negative or whether there were some unknown effects on participants' behavior in this condition. ~~MMI~~ The sensitivity was ~~MMI~~ not affected by either manipulation.

### General Discussion

People often make decisions that incur more risk than is normatively optimal. The research reported here used a novel signal detection theory ~~approach~~ <sup>approach based on</sup> to examine some of the psychological reasons for this phenomenon. ✓

## Summary of Results

shelter → decisions and  
 Using participants' subjective likelihood ratings and decisions for an adverse weather event ~~and our random likelihood model~~, we estimated their subjective criterion for taking a protective action. There were four main results. *spell out*

1) Manipulating both eoc and gain-loss frame significantly affected the subjective criterion but not subjective likelihood. Thus, these results are consistent with the hypothesis of a biased subjective criterion and reject the hypothesis of a biased subjective likelihood.

2) Experiment 1 demonstrated, for that first time, that manipulating the economically optimal criteria revealed a centering effect on the subjective criterion. Specifically, the subjective criterion moved toward 50% relative to the economically optimal criterion. This occurred whether the economically optimal criterion was above or below 50%. ✓

3) In Experiment 2, the manipulation of framing resulted in a shift of the subjective criterion that was consistent with a greater risk-seeking tendency with a loss than a gain. The ~~centering effect was larger than~~ the framing effect *was smaller than the centering effect* in key conditions requiring a protective action *with a* despite the low likelihood of an adverse event (25%).

4) The internal consistency of the proposed random likelihood model was verified in both the reanalysis of previously published experiments and the two new experiments. This supports the key assumptions of the random likelihood model *was supported*. *← Insert 60A*

These results failed to support the biased subjective likelihood hypothesis. Risk seeking choices were not due to the perception that the likelihood of the severe weather events was lower than they were. In fact, and consistent with previous research (Gulacsik et al., 2022; Qin et al., 2024), participants' perception of the likelihood of the severe weather event was reasonably accurate, albeit slightly overestimated in this situation in which probabilistic information is

# GOA-1

Subjective criteria versus subjective outcomes  
Implications for our ~~two~~ two hypotheses } [New Headers]  
^

These results support the subjective criteria hypothesis.

Both of our manipulations affected the subjective criterion and not the subjective likelihood. Furthermore

the for the conditions with low likelihood of an <sup>adverse</sup> weather event, the subjective criterion was

higher than the economy option criterion which resulted in risk seeking choices.

60A-2

These results were similar to  
how the subjective criterion included  
bias in other applications of  
signal detection <sup>theory</sup>, to ~~the~~ decision  
making (Swets et al., 2000).

provided with each forecast. Indeed, the signal detection approach used here revealed that the overestimation of likelihood was counteracted by an even greater bias in subjective criterion.

### The Centering Effect

The results allowed us to identify the role of centering (Poulton, 1979, Olkkonen et al., 2014) on the subjective criterion. It is possible that the centering effect did not shift the subjective criterion to the precise center of the range (50%). Instead, two of three measurements suggest that the shift was slightly lower than 50%. In Experiment 1, the subjective criterion in the 75eoc condition shifted downward to a greater extent than the subjective criterion shifted upward in the 25eoc condition (both still higher than 50%). In Experiment 2, the subjective criteria in both the gain and loss frame 50eoc conditions were significantly lower than 50%, at 43% and 44%, respectively. One possible explanation is that participants' preconceived notion of severe weather events was that the potential harm is much greater than the cost of protection to take action. Therefore, they adopted a criterion that was slightly lower than 50% as such a criterion was consistent with their preconceived notion.

A potential mechanism for the centering effect is anchoring and adjustment (Kahneman, 2003). When people estimate an unknown quantity, they sometimes start with a number recently encountered and then adjust away from that anchor to make their estimate. The adjustment is often insufficient. In the case of the centering effect observed in the experiments reported here, it is possible that the center of the range of possible probabilities (50%) was an anchor from which adjustment was insufficient. Alternatively, it might be the participants intentionally shifted the subjective criterion toward the center of the likelihood range to avoid extreme decisions. This is similar to the compromise effect observed in decision making (Simonson, 1989).

spell out "eoc".

Instr 61A

NLW  
PWT

61A

When the economically optimal  
criterion was 25%, the observed  
subjective criterion was ~~intermediate~~

between 50% and 25%, when

the economically optimal criterion was  
75%, the observed subjective  
criterion was between 50% and 75%.

In short, the subjective criterion was  
shifted toward the center of its  
possible range (50%).

The centering effect did not appear in subjective likelihood in the reanalyses or new experiments reported here. Perceptual studies suggest that when <sup>good</sup> ~~the~~ information regarding the correct response is available, the centering effect is reduced (Radvansky et al., 1995). In the current experiments, reliable probabilistic information was provided in each forecast perhaps overpowering the influence of centering on subjective likelihood. Indeed, in control conditions of previous studies in which probabilities were not provided, effects <sup>consistent with</sup> ~~similar to~~ centering were observed such that the mean likelihood ratings was close to 50% (Demnitz & Joslyn, 2020; Qin et al., 2024). However, when probabilities were provided, the centering effect on subjective likelihood disappeared. ✓

### **The Gain-Loss Framing Effect**

In Experiment 2, framing was manipulated. Consistent with prospect theory (Kahneman & Tversky, 1979), the subjective criterion was higher, that is, more risk-seeking, for a loss compared to a gain. However, with a 25% economically optimal criterion in the gain frame, the subjective criterion was also risk-seeking, although less so than with a loss. Similar patterns were observed in the shift in self-reported criterion and the proportion of trials choosing the safe option, corroborating this analysis. This suggests that a shift in the subjective criterion mediates the risk-seeking and risk-averse tendencies observed here.

However, the effect size of gain-loss framing in this study was small (only 3% difference at 25eoc) in comparison to classic experiments in prospect theory. For example, in the disease problem (Tversky & Kahneman, 1981), in a loss frame a quarter of participants chose the safe over the risky option while in a gain frame three quarters chose the safe over the risky option. There are two potential explanations. First, the multi-trial nature of the experiments reported here might have led participants to consider the big picture as opposed to one trial at a time (narrow

spell out

framing) when making decisions. When considering a series of risky decisions together instead of each trial independently, people are better able to consider the long term strategies that underly an appreciation of expected value (Thaler, 1999) and are less susceptible to the effects of gain-loss framing as a result. An alternative explanation is that in classic prospect theory experiments, the expected values of the safe and risky options were the same or similar. In the experiments reported here, the expected value of the risky option might differ drastically from that of the safe option and varied trial by trial. Therefore, it was possible that the gain-loss framing effect became less salient when the expected values varied widely.

#### Random Likelihood Model

This work introduces <sup>a new</sup> the application of signal detection theory, <sup>the</sup> using a random likelihood model, <sup>It is</sup> based on the assumption that subjective likelihood and a subjective criterion alone determined participants' decisions. To validate this assumption, ROC plots were calculated from likelihood ratings conditionalized by whether or not there was a drought. In 6 out of 7 cases the participants' binary decisions (proportion of hits and false alarms in decisions) were consistent with the ROC curve based on the likelihood ratings. This supports the validity of this approach which isolates the effects of subjective criterion from that of subjective likelihood. In addition, the subjective criterion should remain the same regardless of trial structure, allowing for comparison between studies with different trial compositions.

This novel approach makes important contributions to our understanding of risky decision making. Using a signal detection approach, we identified the subjective criterion and not the subjective likelihood as the mediator of the decision bias in naturalistic decisions. We

[Left-hand side headers]

Fixed 63A

new para

63 A-1

and our application of it

This model is different from previous applications of signal detection theory in two ways. First, it introduces randomness into the ~~the~~ representation of libelous evidence than ability. This is motivated by the court experts wanting libelous from true to true and holding the <sup>cost</sup> ~~ability~~ of alternate choices constant.

Second, rather than assume particular distributions for the noise (e.g. Normal),

63A-2

we use the observed difficulty  
of the liberalized voting to  
estimate the subjective criterion.

This is viable for those  
decision making distributions in  
which participants <sup>found</sup> ~~find~~ it  
relatively easy to report  
aspects of their decision.

also found that both the centering and framing effects modulate the subjective criterion and not the subjective likelihood.

### Potential Behavioral Interventions

An abundance of previous research suggests that people are risk-seeking for losses, especially when the economically optimal criterion is low (Baker, 1995; Joslyn & LeClerc, 2013; Atreya et al., 2015; LeClerc & Joslyn, 2015; Qin et al., 2024). The current experiments revealed that in such situations, risk-seeking decisions are mediated by a higher-than-optimal subjective criterion. In real life situations, a response is often required to protect against adverse weather events that have a low probability of occurrence (low economically optimal criterion) because of the potential for severe consequences. Therefore, one important implication of the findings of the current experiments is that behavioral interventions to lower the subjective criterion might help to reduce risky decisions. Based on the gain-loss framing effect, changing the frame of the situation is a straightforward intervention, but framing might not make sense in some real-world situations (e.g., framing tornadoes as a gain).

On the other hand, there are several possible interventions ~~to address~~ the centering effect. Here, this effect was observed in the subjective criterion but not subjective likelihood for which the objective probability was provided. Therefore, communicating the relevant optimal criterion to participants might reduce the centering effect in subjective criterion as well. For instance, participants might be told "you should choose the safe option when the probability is higher than 25%". The effectiveness of directly supplying both economically optimal criterion information and probabilistic information has not been tested, although some evidence suggests that providing explicit advice to protect whenever the optimal criterion is exceeded improves

*that take advantage of*

*of an adverse weather event,*

decisions, as long as the event probability is also provided (Joslyn & LeClerc, 2012).

Alternatively, it might help to simply bring attention to the criterion by asking participants to estimate the economically optimal probability for taking action when they have sufficient information to do so (Cheng et al., 2012). However this might not be feasible in ~~any~~ <sup>some</sup> real-life scenarios because the potential loss might be difficult to quantify and might differ depending on individual circumstances. ✓

Another possible intervention ~~directed at adjusting~~ <sup>to</sup> participants' subjective criterion is to manipulate what they perceive as the center of the range of possible ~~probabilities~~ <sup>likelihoods</sup>. For example, the forecast could provide a visualized ruler showing a range of ~~probabilities~~ <sup>likelihoods</sup> from 0% to 50% and indicating a specific point as the likelihood of a given weather event, e.g., 15%. This might cause users to constrict their perceived range of possible ~~criteria~~ <sup>subjective</sup> to 0%-50%, making 25% the value toward which estimates would shift ~~through~~ <sup>toward because of</sup> the centering effect. Future studies should test these and other interventions to reduce the subjective criterion and risk-seeking choices.

### Limitations and Future Directions

Theoretically, the placement of the subjective criterion should be based on the expected utility as well as the likelihood of the options. The reanalyses and two new experiments reported here focused on likelihood alone. Consider several examples of this focus. The objective probability of weather events was manipulated from trial to trial while utility was held constant. Participants reported a likelihood rating on every trial but never reported utility. Therefore likelihood ratings but not utility were used to calculate the subjective criterion. Future research might shift the focus to utility to determine the impact on subjective criterion.

With utility in mind, another possible future direction is to incorporate the expected utility (perceived efficacy in risk communications; Rogers, 1975) of protective actions into the model as a variable that can have stochastic distributions. These experiments assumed perfect efficacy of protective actions. This is rarely the case in real world situations. According to protection motivation theory, in addition to one's risk perception (probability and severity of an adverse event) and the perceived efficacy of protective actions is critical to their uptake (Rogers, 1975). In hurricane and flood situations, for instance, perceived efficacy was linked to compliance with evacuation orders (Demuth et al., 2016; Liu et al., 2024). When the efficacy of protective actions is not perfect, people's perceived efficacy of the safe option might vary, affecting the placement of the subjective criterion and as a result, people's decisions. Therefore, it is important to take perceived efficacy into account for a more comprehensive understanding of how people make decisions under risk in anticipation of severe weather events.

Finally, the current random likelihood model operates on the assumption that the subjective criterion has no variability and is fixed across decisions. This is unlikely to be the case in real world situations. Indeed, research suggests that after recent encounters with adverse weather events, decisions became more risk-averse while the subjective likelihood remains unchanged (e.g., Demnitz & Joslyn, 2020). This suggests that the subjective criterion can change between decisions based on recent trial experience. Recent outcomes might not be the only reason for changes in the subjective criterion. The objective probability might serve as an anchor that shifts the subjective criterion from one trial to the next. This ~~was not possible~~ <sup>could not be analyzed</sup> in the current experiments because there were not enough trials to conduct within-group analysis at different objective probability levels. A similar experiment with more trials or a different experimental paradigm <sup>might</sup> ~~could~~ be used to explore this issue. In addition, participants received reliable

probabilistic information in the experiments reported here. Reliable probabilistic information might not be available in many real life situations, and thus these situations should also be tested in future studies. *with unreliable probabilities*

### Constraints on Generality

The new experiments and the reanalyses used an online US nationwide sample that was fairly representative in terms of age, gender, race. However, since the sample was recruited from the Prolific, the education level is expected to be slightly higher than a representative sample (see Prolific, 2025). For that reason, the results might not extend to those with lower levels of education. Nonetheless, as the sample was fairly representative in age, gender, race, so we consider it suitable for generalizing to the US population.

### Conclusion

This study used a random likelihood model based on signal detection theory to better understand the risk-seeking tendency in naturalistic weather decisions with reliable probabilistic information. *The model allows one to distinguish between shifts in subjective criterion and/or subjective likelihood of two old experiments*

The results of the reanalyses and two new experiments suggest that risk seeking is due to shift in participants' *subjective* criterion for taking protective action. Specifically, risk seeking is due to participants requiring a higher likelihood of the adverse event to take protective action than what is optimal according to the *economically optimal* normative standard. In addition, both ~~the~~ gain-loss framing and the center effects were observed in subjective criterion but not in subjective likelihood. For the critical conditions requiring protective action despite the low likelihood of an adverse event, the centering effect in subjective criterion was the primary source of risk seeking.