Dear Dr. Williams and Reviewers,

Thank you for your careful and thoughtful reviews of our manuscript, "The impact of mapper experience and data on the quality of geomorphic fault mapping." We are grateful for the time and expertise both reviewers invested in evaluating our work. Their comments have substantially improved the manuscript, and we hope the revised version addresses their concerns satisfactorily.

Below, we provide a detailed response to each comment from the editor and both reviewers. All changes are highlighted in violet in the marked-up manuscript, and deleted text is shown with strikethrough. Line references in our response text refer to the location in the marked-up manuscript.

The three most substantial changes to the manuscript are as follows. First, we revised our treatment of aleatory variability and epistemic uncertainty throughout the manuscript, adding new paragraphs to the Introduction, Methods, and Discussion that more clearly define and distinguish these two sources of uncertainty, and clarify that our metrics capture an upper bound on epistemic uncertainty due to the unavoidable inclusion of some aleatory variability in our comparison to coseismic ruptures. Second, we reorganized the Methods section to reduce timeline narrative language, consolidating the description of professional geologist participation into the Fault Mapping Course subsection and streamlining participant descriptions to focus on the scientific content rather than course logistics. Third, we added an explicit distinction between mapping accuracy and map quality early in the Methods section, defining accuracy as the spatial correspondence between mapped fault traces and coseismic rupture locations (Metric 1) and quality as the geological reasonableness of mapped fault traces as reflected by the proportion that later ruptured (Metric 2), and reviewed the manuscript for consistent use of these terms.

We have added one coauthor at the revision stage: Benedict Johnson (Ruhr University Bochum), who participated in the study as a graduate student mapper. Dr. Johnson contributed to the Investigation phase of the research, completing the Repeat Mapping Project using the same datasets and instructions as other class participants. He was not included in the original author list as he did not respond to the authorship invitation prior to submission, but has since confirmed his participation and approved inclusion as a coauthor. He was previously acknowledged in the Acknowledgments section.

We look forward to your feedback on the revised manuscript.

Sincerely,

Malinda G. Zuckerman and coauthors

Dear Malinda G. Zuckerman et al:

I hope this email finds you well. I have reached a decision regarding your submission to Seismica, "The Impact of Mapper Experience and Data on the Quality of Geomorphic Fault Mapping ". Thank you once again for submitting your work to Seismica.

Based on reviews I have received, your manuscript may be suitable for publication after some revisions.

I have now received two reviews for your manuscript. Both reviewers are generally enthusiastic about the work. Both reviewers also suggest, however, that revisions will be required to bring it into a more acceptable state for publication. The most common criticism raised by both reviews concerns the usage of epistemic versus aleotoric sources of uncertainty. Although this might be considered an issue of lesser salience in many contributions, in this case these ideas are central the overall message of the manuscript. Reviewer #2 also notes the relevance of this work to seismic hazard analysis, specifically probabilistic fault displacement hazard analysis, where these are first-order terms of practical importance for model construction. Although correction of these issues may constitute a volumetrically minor component of your revisions in the eventual product, I would encourage you to undertake them with some detailed consideration. For example, the reviewers point out several instances where the categorization of uncertainty as epistemic or aleotoric may not be so simple, and perhaps is more a function of what goal the data are intended to achieve.

A point of discrepancy between the two reviewers concerns the writing clarity and/or presentation quality. The first reviewer comments that they feel these are in reasonable state in current form. The second reviewer notes some issues with manuscript organization that detract from overall clarity. For my part, I tend to side more with the second reviewer. There are some grammatical and/or punctuation issues throughout the manuscript. The second reviewer provided many detailed comments and suggested edits to address these in a pdf that I will link to you in the editorial portal. They also note that the organization is perhaps a bit cumbersone. Having read through the manuscript again after receiving the reviewer comments, I suspect the organizational concerns may stem in part from a slightly mixed approach to manuscript construction. Insofar as I understand the intent of this manuscript, it is to focus on a fundamental issue within active faulting and seismic hazard studies, specifically the role of experience in prediction accuracy and/or uncertainty. In various instances, however, the manuscript reads a bit more like an educational study, where specific student activities, coursework, and timelines are described in a detailed narrative. I do of course understand that some of the primary data collection for this work was effectively conducted by students in a coursework environment. I do think though that the organization might be improved by constructing a more traditional (i.e. less timeline narrative focused) methods

section that links forward to results and discussion. Please understand that this is only my suggestion. Other routes toward improve organization and clarity may also be valid as you look toward revisions.

Please do not hesitate to reach out if you have questions, and thank you for considering Seismica for publication of your work.

When you are ready to resubmit the revised version of your manuscript, please upload:

- A 'cleaned' version of the revised manuscript, without any markup/changes highlighted.
- A pdf version of the revised manuscript clearly highlighting changes/markup/edits.
- A 'response-to-reviewers' letter that shows your response to each of the reviewers' points, together with a summary of the resulting changes made to the manuscript.

Once I have read your revised manuscript and rebuttal, I will then decide whether the manuscript either needs to be sent to reviewers again, requires further minor changes, or can be accepted.

If you deem it appropriate, please check that the revised version of your manuscript recognises the work of the reviewers in the Acknowledgements section.

Please note that Seismica does not have any strict deadlines for submitting revisions, but naturally, it is likely to be in your best interest to submit these fairly promptly, and please let me know of any expected delays.

I wish you the best with working on the revisions. Please don't hesitate to contact me with any questions or comments about your submission, or if you have any feedback about your experience with Seismica.

Kind regards,
Randy Williams

We thank the editor for his observation regarding the methods section's timeline and agree that it benefited from a more traditional structure. We made several targeted revisions to reduce the timeline narrative and focus on the scientific content. Specifically, we consolidated the description of professional geologist participation into the Fault Mapping Course subsection (Section 2.2) rather than presenting it as a separate subsection, removed several sentences describing course logistics (meeting schedule, lead author's student status), and streamlined the participant description to focus on the mapping activities rather than the course timeline. We believe these changes bring the methods section closer to a traditional format while preserving the context necessary to understand the study design and the sources of the data.

Reviewer 1:

This is a review of “The impact of mapper experience and data on the quality of geomorphic fault mapping” by Zuckerman and coauthors. This study analyzes and compares fault mapping from a group of geologists with varying levels of experience, using datasets of different resolutions and different case studies of recent surface ruptures. By comparing a dataset of pre-rupture mapping to surface traces of recent surface ruptures in a variety of settings, the authors conclude that fault mapping experience does not significantly improve mapping accuracy. Higher-resolution data and regional geologic context, along with a systematic approach to mapping, typically yields higher quality mapping and accuracy. The manuscript is well written and illustrated, and is logically organized so that the reader can easily follow the different subsets of data and their results. The interpretations provide a quantitative epistemic uncertainty on fault-location mapping, and overall, the interpretations and conclusions in this study will be of interest to many researchers in the earthquake geology, paleoseismology, and hazard mapping communities. This is a well-thought out and very interesting study on how we map faults and their uncertainty, and how it may vary with mapper experience, data resolution, and additional context. However, I think there are several limitations that should at least be mentioned, perhaps in the discussion, introduction, or methods, to help the reader better grasp some of the background and assumptions of the study and place the variables studied here into better context. I mention these in my broader comments below, followed by minor comments by line number, and minimal comments on the already wonderful figures. I look forward to seeing this work published!

Major comments:

1. Were the mapped fault surface rupture traces used for comparison also mapped remotely, or were they mapped in the field? I did not see this clearly stated in the text (although perhaps I missed it), and many of the cited studies for each case study include a mix of both remote and field mapping. The reason I think this is an important issue to mention is that the resolution of what we can be seen in remote sensing datasets pre-rupture versus detailed field post-rupture mapping varies from site to site, and it may not necessarily be fair to state that the remote mapping didn’t align with the most recent coseismic rupture mapping if coseismic ruptures are distributed and small in the most recent event, or the fault has a history of this rupture character such that it may not leave a clear landscape record. This topic is alluded to several times throughout the manuscript, but the reader would benefit from some clarity on this limitation and why the comparisons are made using different resolutions (or not) in different case studies.
   This is a valid point. We clarified how the coseismic rupture traces were mapped in

Section 2.5 (Lines 215-7), because this is where we first mention the ruptures. The Borah Peak ruptures were mapped in the field, while the El Mayor-Cucapah, Fukushima-Hamadori, and Ridgecrest ruptures were mapped using a combination of field mapping and remote sensing data. We added a parenthetical note with the relevant citations at the point where we first introduce the published coseismic rupture traces in the data analysis section.

2. I think it's implied in the definition of epistemic and aleatory variability in the introduction, but a major assumption in this study is that the fault has ruptured in a similar fashion in the geologically recent past, such that the landscape is an appropriate and accurate recorder of past surface faulting events. It would benefit the reader to have this assumption more clearly stated up front and discussed as a limitation of the study. Note that this is nicely illustrated by the El Mayor-Cucapah part of the study, where two ruptures occurred with overlapping rupture segments, but faults closer to the range did not clearly rupture in the last 2 earthquakes. I see this is briefly mentioned in L485-487, but could be expanded. I also realize this is a not a key point of this study as it covers more of the aleatory variability aspect, but I think a more thorough discussion of this would strengthen the manuscript.
   We agree that the assumption underlying geomorphic fault mapping — that past ruptures have left a detectable landscape record — should be explicitly stated. We added a paragraph to the Discussion overview (Section 4.1; Lines 487-491) noting that the clarity of the geomorphic record depends on rupture recency, frequency, and the rate of subsequent landscape modification, and that where this record is incomplete, mapping accuracy is limited regardless of experience or data quality.

3. The reader may also benefit from some context on the recurrence interval or most recent event on each fault. For example, if a fault had ruptured more frequently, or had an MRE more recently, it may have a stronger landscape record (as demonstrated by the El Mayor Cucapah case study, and thus may have been more easily mapped by mappers of all experience levels. Although briefly mentioned in the introduction, this caveat on variable landscape records from past fault activity may also influence the subtleness and hence ability of mappers to recognize fault-related topography. I also see this is briefly mentioned in L522-524, but could be expanded.
   We agree that recurrence interval and most recent event context would help readers interpret the variable mapping difficulty across our study areas. We added a clause in Section 2.3.1 (Lines 161-2) noting that the clarity of tectonic landforms reflects a combination of climate, fault slip rate, and time since the most recent surface rupture, citing Scott et al. (2024) which directly addresses how these factors influence fault displacement hazard uncertainty.

Minor Comments:

L67-71: This paragraph seems a little out of place here… consider moving it or weaving these definitions into the following paragraph for flow.
We agree that the original paragraph was out of place. We replaced it with a new paragraph (Lines 70-6) that introduces both aleatory variability and epistemic uncertainty as two sources of uncertainty in predicting future rupture locations, providing a natural transition from the discussion of fault mapping variability to the literature review. The cognitive bias sentence was removed as it was tangential to the definitions, and "aleatoric" was corrected to "aleatory" for consistency with standard seismic hazard terminology.

L90: Consider changing to “This study”, because the authors also conducted the two studies mentioned in the previous paragraph.
L105: We made this change.

L138-143: These specific challenges for each earthquake are hard to place into context without the background knowledge of each event. I realize these are described in detail in subsequent subsections, but consider at least adding the fault slip sense to these brief descriptions up front, to help provide initial background and better illustrate the differences.
L163, 164, 166: We added the fault slip sense.

L184: Consider modifying this statement to “appear to host less or no slip”, because a subtle landscape record may also be affected by preservation of the landform, right?
L211: Yes, that is a good distinction. We made this change.

L232: Define Theta in equation 2.
L280: We added “θ” to the sentence before the equation.

L431: Should this be Table 1?
L496: Yes, thank you, this is a typo. We made this change.

L448: Consider citing some of the figures here that demonstrate a similar quality.
L521: We added references to Figures 4, 5, and 6 to support the claim that graduate students and professional geologists produced maps of similar quality.

L461: Consider clarifying to “similar fault systems or settings” instead of “projects”.
L537: We made this change.

L510-512: I appreciate this statement from the participants! I often find that really high resolution is not always better – sometimes the addition of too much noise masks the tectonic signal.
Yes, this was an interesting finding!

L561: Could “modestly better” be quantified?

L656-7: We replaced "modestly better" with a more precise description of the performance difference, noting that professionals performed somewhat better than undergraduates but comparably to graduate students, with a reference to the supplemental table.

L593: It may help the reader here to compare how these widths compare to existing fault zone widths used in PFDHA from Table 1.

L698-700: We added a sentence comparing our epistemic uncertainty values (2--133 m) to the regulatory fault zone widths in Table 1, noting that they span from below the Alquist-Priolo setback distance at well-defined scarps to approaching the regulatory fault zone width at ambiguous landforms. This directly addresses the reviewer's suggestion to contextualize our results within the existing PFDHA framework.

Figure 1: Consider adding a brief label or description of each sub-figure to the caption, and perhaps adding sub-figure labels.

We believe Figure 1 is sufficiently self-explanatory. All locations are labeled by name, countries and states are identified, and latitude/longitude coordinates are provided on each panel.

Figure 4: This figure is very helpful to understand the results of the mapping. I wonder if it’s possible to change the heatmap color ramp? The Min (white) gets easily lost in the white-gray of the shaded relief map. (Note that in Figure 5, the shaded relief map is more gray, so the white Min stands out a bit more.

We appreciate this suggestion. However, the hillshade basemap for Borah Peak is derived from a 1966 digital surface model with blocky artifacts (Reitman and Briggs, 2024), and adjusting the hillshade contrast to improve the color ramp visibility would make these artifacts more prominent and potentially misleading. We therefore retained the original color ramp.

Figures in the supplement may benefit from lat/long coordinates to help the reader better place the location context.

We added approximate center coordinates to the supplement figure captions to help readers place the location context.

Recommendation: Revisions Required

Reviewer 2:

I think this is an interesting study that addresses several important concepts in seismic hazard studies. They have clear, applied implications for probabilistic fault displacement hazard studies, and to a lesser degree, probabilistic seismic hazard studies as well. I would strongly encourage you to continue exploring your data and results. Two broad comments are provided below, and the more specific comments in the attached document. I also encourage you revise your approach to organizing the manuscript and improve the writing in several areas. After doing so, I think this would be a valuable contribution to applied seismic hazard studies.

I have two fundamental comments that I believe should be addressed/explored.

I believe the distinction between aleatory and epistemic uncertainty should be examined more carefully. In several parts of the paper, aleatory variability appears to be inadvertently treated as epistemic uncertainty. For instance, there is little discussion of the aleatory component associated with which strand of a fault may or may not rupture during an event, or when a new strand develops and rupture to the surface. When rupture occurs on a strand that shows no evidence of past surface rupture but develops in the most recent event, it is unclear how this is represented in your framework. I would argue that such cases are typically captured within the aleatory variability of the rupture model—reflecting the inherently random nature of rupture extent—rather than as epistemic uncertainty. This is sort of a surface effect, driven in part the material properties that rupture passes through. In other words, even with additional data, the occurrence of rupture on that unrecognized or previously non-existent strand could not have been predicted in an informed manner. These situations likely fall outside the scope of your current treatment of epistemic uncertainty and should be evaluated separately. I believe most fault rupture models account for off-fault rupture through aleatory terms.

We thank the reviewer for highlighting the importance of clearly distinguishing between epistemic uncertainty and aleatory variability in fault mapping and hazard analyses. We agree that our original framing could better articulate how these two components interact in our study. Our experiment primarily quantifies epistemic uncertainty, defined here as variability in mapped fault location arising from differences in mapper interpretation, data quality, and dataset resolution. However, we acknowledge that our comparison to real coseismic ruptures necessarily includes a component of aleatory variability—that is, the inherently random spatial variability of surface rupture location between successive events on the same fault.

To address this comment, we have:

1. Added a new paragraph in the Introduction (Lines 70-6) that defines both epistemic and aleatory uncertainty in the context of geomorphic fault mapping and probabilistic fault displacement hazard analysis (PFDHA).
2. Revised the Methods (Section 2.5, Lines 238-40) to explicitly note where aleatory variability is implicitly captured in our metrics (e.g., rupture-to-fault separation distances) and to clarify that our uncertainty values represent an upper bound on epistemic uncertainty.
3. Added a paragraph in Discussion 4.5 (Lines 686-94) that explains how our framework can complement PFDHA approaches, where aleatory and epistemic terms are typically parameterized separately.

These revisions clarify that while our repeat mapping experiment quantifies interpretation-driven uncertainty (epistemic), the real-earth comparison unavoidably incorporates some aleatory variability associated with rupture segmentation, strand activation, and surface expression.

Aleatory variability vs epistemic uncertainty concept ties in to my other fundamental comment, which is related to the mapping, and finding the "correct" answer. For example, if someone predicts a surface rupture along a strand showing clear evidence of past ruptures, but the MRE does not actually rupture that particularly strand, how is that situation managed in the data analysis and results? It seems to me that the mapper would have been justified to predict that section would re-rupture based on the tectonic geomorphic landscape; however, if I am following your approach correctly, it seems that is treated as a "miss". Similarly, when a rupture occurs at a surface location that had never ruptured before, then was the practitioner somehow wrong when they missed it? This of course ties back into my first fundamental comment on variability and uncertainty.

We appreciate this thoughtful comment regarding how our analysis defines "accuracy" relative to the observed coseismic ruptures. We agree that the concept of a "correct" answer in fault mapping is nuanced. In our framework, accuracy is operationally defined as the spatial correspondence between mapped pre-rupture fault traces and the coseismic rupture location from the next event. This definition provides a reproducible benchmark for quantifying mapping performance but necessarily treats the observed rupture as a reference realization of rupture location, not as the sole "truth."

We acknowledge that in cases where a mapper reasonably identified a geomorphically expressed fault that did not rupture in the subsequent event, their interpretation is still geologically valid. Such cases reflect aleatory variability, not mapping error.

We have clarified this point in the Methods (Section 2.5, Lines 229-234) and Discussion (Section 4.3.2, Lines 574-8), noting that our "misses" may represent valid fault interpretations that simply did not rupture in the studied event. We added one paragraph in each section.

Recommendation: Resubmit for Review

Line-by-line comments:

29: Perhaps since however was just used, consider an alternative sentence structure or word choice, such as although, in contrast, etc. Or replacing the prior however.
L30: We replaced this second however with “in contrast”

30-31: this sentence is a bit awkward. Can it be simplified a bit, e.g. to test the efficacy of mapping experience on mapping quality, …
L32: We agree that this sentence reads a bit heavy, but we wish to keep the full meaning of the sentence. We did, however, replace “how much” to “to what extent” to read more formally.

35: Clarification needed: e.g. … missed identifying tectonic landforms indicative of future ruptures
L37: We added this clarification verbatim.

36: I think this needs context. I assume this is the epistemic uncertainty in the fault's actual location, or does this reflect the inter-mapper uncertainty?
L38: We changed the sentence to, “Using participant mapped fault data, we quantify inter-mapper epistemic uncertainty in fault location to be between 55 and 117 m.” We believe adding these 4 words adds context while keeping the abstract brief.

37: just my opinion - this is more of a suggestion than an indication.
L39: We changed “indicates” to suggests.

55a: characterizing
L57: We made this change.

55b: evaluating or assessing
L57: We chose to use “evaluating”

58: not sure why his would be "potentially" if the sentence goes on to mention in a definitive sense the landforms are generated by repeated ruptures. I think either delete potentially, or leave it and modify the rest fo the sentence to indicate the landforms are interpreted as being tectonic.
L60: We appreciate this point. We deleted “potentially”

60a: complicated
L62: We made this change.

60b: i think you are really referring to interpretability rather than visibility here?
L63: Yes, interpretability is more specific. We made this change.

62-63: I think you are referring more to geomorphic mapping of faults than fault mapping from a structural geology perspective.
L66-7: Yes, the participants mapped using geomorphology, but understanding common fault expressions from structural geology knowledge helps. We replaced this sentence with "Geomorphic fault mapping is inherently interpretive, requiring geomorphic observation and benefiting from structural geology training to identify and map tectonic landforms."

67: Since you are really introducing concepts of uncertainty and variability, I think you need an introductory sentence here, e.g. "uncertainty in predicting the location of future surface ruptures comes from two sources: aleatory variability, or the inherent randomness of process, and epistemic uncertainty, which reflects limitations in current knowledge."
L70: We added this suggested transition. It leads in well to the edited paragraph.

69: not sure this is necessary. Cognitive bias also stems from a lack of training or knowledge.
We removed this sentence when revising this paragraph.

70: since this is ultimately a seismic haz audience, I suggest aleatory.
L71: Thank you for catching this mistake. We made this change.

72: needs a stronger intro sentence with more context.
L82: We replaced this sentence with, "Uncertainty in mapped fault locations has long been recognized as a limitation in both tectonic geomorphology and fault hazard studies" to add more context and make it clear that this is the literature review.

72-73: not entirely sure what this means. Location error of faults relative to what? It sounds like there is a measured location of an actual ruture to the pre-event mapped trace?
L84: We added "mapped" before 'fault location' to show that the location error is between the mapped faults and the rupture traces.

75a: specify the type.
L86: We clarified that Thompson Jobe et al. (2020) used a combination of pre-event remote sensing data and field verification. The study incorporated numerous remote sensing data types at varying resolutions, which we did not enumerate in full to keep the introduction concise.

75b: missing period
L87: We made this change.

77-78: I'm not sure what this means. I think you are talking about different mappers interpreting the same features (pre earthquake) differently?
L90-1: Our sentence was unclear. In the study, different mappers, using post-earthquake data, interpreted the same features differently, causing a difference in their rupture mapping. We changed this sentence to, “Young et al. found that mappers using post-earthquake remote sensing data produced rupture maps that differed by 10-20% due to interpretive differences in geomorphic expression. Additionally, map…” and onto the next comment.

78: would accuracy be a better term? Quality would suggest that when fatures are very subtle, good quality cannot be achieved.
L92: Yes, we agree that accuracy would be better. We made this change.

80-81: unless this reference has context that is described somewhere, this seems unnecessary
We retained this reference as Tikoff et al. (2023) is discussed in the context of our results in Section 4.4.2, where we describe the benefit of assigning confidence ratings to fault traces.

81-83: the first 3/4 of paragraphs gets into background of studies that dealt with mapping accuracy. This introduces a somewhat different set of concepts that I think woudl benefit from it's own paragraph to elaborate on it (if it is supportive of your general thesis). Alternatively, you could perhaps weave it into this paragraph with clear context that relates it to the rest.
We chose to retain this content in the same paragraph rather than creating a separate paragraph, as we feel it flows naturally as a concluding point to the literature review on fault mapping uncertainty. The bias discussion is expanded in Section 4.3.1 where it is interpreted in the context of our results.

84: this sentence shoudl be strengthened. "...that compared pre-earthquake mapped fault traces with coseismic ruptures, and addressees how standardized mappig approaches can improve fault mapping". and then go onto elaborate a bit on the two studies.
L97-8: We made this change to the topic sentence. We left the rest of the paragraph the same to save on word count.

85a: I think this is pre-event, mapped fault traces?
Yes. With the additional text in the previous sentence, this is now more apparent.

85b: from the _____ earthquake(s)
L99-100: In lieu of listing all 7 earthquakes, we added “from seven M 6.0-7.8 earthquakes.”

87: "less experienced" or "novice" may be a better phrase than 'junior'
L101: We agree. We changed “junior” to “novice”

88: what is "the" standardized mapping approach? "a" standardized…?
L103: We added “introduced in Scott et al. (2023)” to add context.

90a: the level of training/experience; 90b: remove “level”; 91: Change “location” to “mapping.” Are you evaluating mapping accuracy, uncertainty, or both?
L105-7: We agree with these suggestions. We changed "Our" to "This" to avoid ambiguity and follow the suggestion from Reviewer 1, restructured the phrase "mapper experience level" to "the level of experience," and revised the final clause from "fault location uncertainty" to "mapping outcomes and epistemic uncertainty" to more accurately reflect that the study evaluates both accuracy and uncertainty. The revised sentence reads: "This study quantitatively evaluates how the level of experience affects fault map accuracy and offers insights into how training, experience, and a standardized mapping approach impact mapping outcomes and epistemic uncertainty."

94: who all used a standardized mapping approach
L108-11: We changed the sentence to, “We assessed and compared maps created by 23 mappers, ranging from senior undergraduate students to senior professional geologists, who all applied the same standardized, structured mapping approach.” to make it explicit that everyone used it.

95: While you are evaluating epistemic uncertainty, aleatory variability is also a part of this since surface ruptures can vary at the surface. so while a mapper may be spot-on mapping prior ruptures, the next rupture could be in a different location due to aleatory variability.
We agree that aleatory variability is embedded in our comparison to coseismic ruptures. We address this throughout the revised manuscript, including in the new paragraphs in the Introduction (Section 1; L70), the Methods (Section 2.5; L235), and the Discussion (Section 4.5; L686), which clarify that our metrics capture an upper bound on epistemic uncertainty as they incorporate some aleatory variability in rupture location.

98-100: seems this could be reserved for the conclusion.
L117-8: We removed this sentence and “however” in the next sentence to agree with the change.

101: same comment as above: you are also accounting for the aleatory variability of rupture location.
We note that the epistemic uncertainty quantified from the repeat mapping around tectonic landforms is more purely epistemic than Metrics 1 and 2, as it measures the spread of mapped fault traces across mappers relative to a common landform rather than comparing to actual coseismic ruptures. The aleatory component is therefore minimal in this analysis. We clarified this distinction in the revised manuscript through the new aleatory/epistemic paragraphs added in response to Reviewer 2's fundamental comments.

117: comprising
L135: We made this change.

118: Delete “Lead author Zuckerman was a graduate student”
L136-7: We made this change.

119: participants
We did not make this change. In the manuscript, we use "participants" to refer to all individuals in the study, including both students and professional geologists. The sentence at L137 describes only the class composition.

127: , which was
L145: We made this change.

132-133: seems like you said this int he prior sentence
L154-5: Agreed. We made this change.

139: think you need to expand on this. A step-over with alluvial fans is not really a challenge in and of itself. they are more like attributes.
L163-4: We replaced this challenge with what we think is the real issue, that there are unexpected ruptures outboard of the range front. Borah Peak “ruptured, in some cases, oblique to the range front”

140: maybe "arid" is better?
We chose “desert” to be consistent with the Köppen Climate Classification scheme.

142a: arid
See previous comment.

142b: low slip on each strand, but is the cumulative fault necessarily low?
Yes, the cumulative geologic offset for Ridgecrest is less than 1 km, compared to EMC (20 km) and Fukushima Hamadori (10 km).

144-145: vaguely familar with the earthquake, or the region?
L170: We added “the 2019 Ridgecrest earthquake sequence” to be more specific.

145: I think if the data are integral to the study, they should appear in the manuscript rather than the supplement.
We chose to retain this table in the supplement for two reasons: (1) Only one participant had exceptional familiarity with one mapping area, and (2) Seismica asks that we include as few tables as possible, preferably only one. If more participants were intimately familiar with the mapping areas, then it would be important to include in the main text.

150: used in this study
L175: We made this change.

152: you give the "name brand" of the 30 m data, but not the others. I suggest being consistent and providing name branding for each. Also, if the high res top data is derived from lidar, state that. Or SFM.
L177-8: We changed this part of the sentence to “and high-resolution airborne lidar- and/or Structure-from-Motion-derived topography datasets (1 m/pixel or less)” to be more specific.

155-156: you mentioned this in previous section
L181-2: We deleted this sentence.

161: published or existing
L185: We made this change.

163a: Suggest editing since the kinematics are ot mentioned for the Napa event. Alternatively, leave it but mention for Napa as well.
L188: We made this change.

163b: previously mapped at 1:10,000 scale
L189: We changed this part of the sentence to “…and for which a published 1:10,000-scale geologic map was available.” We wanted to keep your suggestion while making it explicit that the geologic map was available and not make it read as if only the ruptures were published.

167: I think this section would benefit (or describe earleir in manuscript) the rationale behind including this group, who lacked the course that "baselines" the grad and undergrads. I think it is important to elaborate that point so that you can justify comparison.
We moved the information about professional geologists up to the Fault Mapping Course subsection earlier in the data and methods and added that the professionals’ data broadens the range of experience level (Lines 147-150).

168a: their?; 168b: Remove “level”
L147: We changed this sentence to, “To broaden the range of mapper experience in the study, we also included six professional geologists.” It was previously unclear.

170: this is listed here but not for the other groups (students, grad students, etc). Suggesting more consistency, or a general statement that applies to all, e.g. all participants, regardless of expereience completed the course in 40 hrs…
We addressed this comment. Since this information is now earlier in the Data and Methods section, following the description of what the students did, it makes more sense that the students in the class had half a semester, and the professionals had 40 hours.

177a: fault traces are really a type of tectonic landform. Can you include this in the prior category? Also, they are really mapping lineaments that they interpret as fault traces, right? and landforms that they interpret as tectonic?; 177b: ranks features based on their degree of confidence they are tectonic?
L203-4: We agree with the reviewer's point that mappers are interpreting lineaments as fault traces and assessing confidence in their tectonic origin. We added clarifying language to steps (4) and (5) to reflect this. We kept the overall structure of the workflow consistent with the published methodology in Adam et al. (2025) and Scott et al. (2023) to avoid inconsistency across the related papers.

180: what they interpreted as tectonic landforms
L207-8: We made this change.

181: points of reference
L208: We made this change.

183: I think a table or figure that explains this fault mapping approach would be helpful to the reader. Perhaps remove that "mapping process" from the next figure and create it's own figure with more detail.
Figure 2 is dedicated to illustrating the mapping approach, including both a flow diagram of the workflow (panel A) and a step-by-step example of one participant's mapping for Fukushima-Hamadori (panels B-E). We believe this figure already provides the visual explanation of the mapping process that the reviewer suggests, and we therefore did not create a separate figure.

188-190: I think this would benefit from more explanation. They read the same to me. Both seem to be the percentage of predicted ruptures from pre-rupture mapping but I also understand what you are trying to do. Can you revise this to clarify?
L219-20: We agree that the two metrics were difficult to distinguish. We restructured the sentence to use parallel construction, making clear that Metric 1 is the percentage of coseismic ruptures predicted by the pre-rupture fault maps, and Metric 2 is the percentage of pre-rupture mapped faults that later ruptured.

190-192: based on length? number of sections? Also, worth noting what the published results you compared them to (presumably the correct answer?) were based on, e.g. rapid response lidar, ground mapping, photo mapping, etc.
The length-based calculation is described in detail in step 3 of the methods, following soon after this sentence, so we did not add a forward reference here. The nature of the published coseismic rupture traces is addressed in Section 2.3.1, Project 1: Repeat Mapping (mapped in the field following each earthquake, with additional remote sensing data used for El Mayor-

Cucapah, Fukushima-Hamadori, and Ridgecrest) in response to Reviewer 1's Major comment 1. (Lines 215-7)

194: eary on I made a comment about "quality". I am still not clear how quality is determined. I can guess, but I think it is worth spelling out, explicitly (early on).
L224-8 and throughout: We agree that the distinction between accuracy and quality needed to be made explicit early in the methods. We added a sentence defining accuracy as the spatial correspondence between mapped fault traces and coseismic rupture locations (Metric 1), and quality as the geological reasonableness of mapped fault traces as reflected by the proportion that later ruptured (Metric 2). We reviewed the manuscript to ensure these terms are used consistently, noting that broader uses of "quality" in the discussion refer to the overall standard of a map encompassing both metrics rather than Metric 2 specifically.

195: can you be more specific about whether the buffers surround the interpretation of the participants, of if they surround the post-earthquake actual ruptures?
We believe that this information is written explicitly in the following two sentences.

199: should be in front of prior figure.
The paper will be reformatted into two columns upon publishing, so all text will change position.

200: i commented in the capturing of not only epistemic, but also aleatory varibility. I think to some degree, the buffer zones account for some of it, at least the site effect that controls the vaibility in the actual rupture location. But it still does not account for things like rupture length since magnitudes and thickness exert some control on those things (amongst others).
L251: We revised the text to explicitly state that the buffer widths capture spatial variability in rupture location relative to mapped fault traces, but do not model other sources of aleatory variability such as rupture length or magnitude-dependent rupture extent. The text now reads, “This ensemble represents the distribution of coseismic surface deformation **with respect to fault-zone width and rupture location relative to mapped fault traces**.”

205: so this half width or buffer is sort of an aleatory treatment?
L255-7: We clarify that while the buffer widths are informed by observed rupture distributions and therefore include some aleatory variability, they are not intended as a formal aleatory treatment. Instead, they provide a practical spatial tolerance for evaluating interpretation-based (epistemic) differences among mapped fault traces. We added the sentence, “These buffer widths are informed by observed rupture distributions and therefore include some aleatory variability, but they are used here as a practical spatial tolerance for comparing mapped fault locations” in the next paragraph.

209: for what? Building codes for commercial and residential?
L258: We clarified that the 15 m buffer corresponds to regulatory fault-zone setbacks “for structures intended for human occupancy”, consistent with California Geological Survey (2018) guidance.

220: In general, this approach all seems valid but I think it woudl benefit form more sytematic organization. Each of your numbered sections starts differently, e.g. this one begins with the main topic and a colon, whereas the preceeding two describe the first (or only) step in the method/calculation
L270: We revised step 3 to use active voice, consistent with the structure of steps 1 and 2.

250-251: are some of these distributed ruptures actually first-time ruptures? or is ther a basis to say they are in fact mappable features?
L301: We clarified that the distributed ruptures were missed in part because their subtler geomorphic expression degrades more rapidly than primary fault scarps, citing Zuckerman et al. This distinguishes the challenge of mapping secondary faults from mapper error or aleatory variability.

258-259: I suggest making this a standard sentence rather than within parentheses.
L309-10: We made this change.

267: this referencing is perhaps misplaced. The 2010 earthquake itself does not need a ref. The ref can go wherever there is context for it (i.e what specifically is Samiento getting credit for in this case?)
L319-21: We reorganized the sentence to have the refs referring to the rupture data.

283-284: I think this is an example of where the aleatory piece plays a role. If these were accurate mapped, but did not rupture, is there something learned from that? There does seem to be a random compoentn to mapping 10+ yr old ruptures, where the sesimogenic fault re-ruputres but the surface traces may not.
L337-9: We agree that mapping the range-front faults was geologically reasonable given their clear geomorphic expression, and that their absence from the 2010 rupture reflects aleatory variability in rupture strand activation rather than mapping error. We added a sentence acknowledging this distinction.

289: THis currently implies that all evidence is obscured. Woudl it be more accurate to say "most"?
L344: We added “may” to show that there is the possibility of evidence becoming obscured.

302: general comment: regarding predicting the rupture locations, it seems that a heavy-handed lineament mapper (i.e. just map everything even remotely resembling a fault scarp"

will always generate the highest % of predicting ruptures. A more experienced person who is more careful with their linework may miss some. I am wondering if there are any efforts to account for this sort of thing.
We agree that a mapper who traces all linear features may predict more ruptures at the expense of overmapping. We added text to the Discussion (4.2 What defines a high-quality map?; Lines 509-515), clarifying that our two complementary metrics capture this tradeoff by balancing rupture prediction with the proportion of mapped faults that later rupture, and that high-quality fault maps reflect a balance between coverage and interpretive restraint.

308-309: if this is your result, i would modify to clarify that and add something like "consistent with" Scott et al., (2024).
L364: We made this change.

313-314: I suggest deleting becasue arid lands mapping is regarded by many as the easiest to map due to the lack of vegetative cover and generally slower geomrophic processes. Also, I don;t think it really adds anything to the section.
L368-9: We removed this sentence.

325: what does this last part actually mean? It seems rather vague. Is it necessary here?
L379-80: We removed the speculative clause at the end.

331: Isn;t this generally associated with SS more than N faulting?
L386: Yes, the way that was written was confusing. We removed that list item, “long linear fault”.

331-333: THis needs a bit more. I think the latter part holds up if you state that the participant was a senior undergrad or someone lacking experience.
L387: We clarified the text to specify that this observation was made by a professional geologist, removing ambiguity about the participant’s experience level.

334-335: this sort of contradicts the last sentence in the prior paragraph .
L390-1: We added some text to separate interpretive differences and spatial prediction accuracy between experience levels. “Although professional geologists differed in their interpretations of slip sense and data quality, these differences did not translate into more accurate predictions of the subsequent rupture location.”

336: Do you mean that *interpreting* a SS at the base would be justifiable?
L392-3: Yes. We changed the text to “consistent with a plausible strike-slip interpretation”.

338: these regular references to the scott paper are providing important context, but they often stand alone and are made without being woven in to the general section's theme. A simple addition of "consistent with Scott et al., " or "Unlike the results from Scott et al., " would

help the reader differentiate this paper's results from other studies.
L395-6: To add more context, we removed the last sentence of the section and used that idea as an introduction to the Scott et al. reference. The paragraph now begins with, “These results are consistent with prior observations that experienced mappers can disagree when interpreting similar geomorphic landforms.”

351a: clarify what this means. Base of ridge, along crest of ridge, etc.
L409: We added that the participants mapped along the “base of the ridge”.

351b: delete “ultimately.”
L409: We made this change.

352: similar comment: I assume this is along the base of the vertical slope?
Graduate students mapped fault traces along much of the slope face rather than only at the base, which is shown in the accompanying figure.

361: assuing you are referring to the systematice approach, I think there are better words than set-up.
L419: We revised "set-up" to "sequential dataset design" to more precisely convey that this section follows the Project 2 design in which participants mapped using each dataset in sequence. It does not refer to the systemic mapping approach.

367: since there is no real context to this, point to the figure that shows its location. Ideally, indicate where it is on Fig 8a-f
L425: We added this in parentheses at the end of the sentence: “(center of Figures 8A–E)”.

371: do you mean "their interpretation confidence"?
L429: We added “ranking” to clarify that the mapper changed the mapping visualization to convey their interpretation confidence. This language, fault confidence ranking, is consistent with how we described the fault mapping approach earlier in the manuscript.

372: what does this refer to? Multiple generations of imagery?
We retained "successive datasets" as this correctly refers to the sequential order in which participants mapped each dataset, which is a key feature of the Project 2 design described in Section 2.3.2 and Table S3.

374: I am getting a bit confused by 1m/pixel sat imagery, and the 1m/pixel airphotos you mentioned earlier. Based on this sentence they are two different things? Looking back at text, I am further confused by it since ther are places that refer to 1m satellite imagery / air photo. I suggest clarifying in a table.
L432-3: We agree that the distinction between dataset types would benefit from clarification. We added descriptors to distinguish the 1 m/pixel color satellite imagery from the 1 m/pixel

black-and-white NAPP airphoto in the text. A full description of all datasets used in this project, including their types and resolutions, is provided in Table S3 in the supplemental materials. Seismica requests that tables be limited in the main manuscript, so we have kept this information in the supplement.

375: mapped
L444: We made this change.

383-384: I think think is worth exploring a bit in your discussion.
We agree this is an interesting observation. However, as this reflects the behavior of a single participant, we do not have sufficient data to draw conclusions about the cause. We have retained the original text noting that the reason is unclear rather than speculating about the participant's decision-making process.

389: expand to clarify. Early you say SFM results in the most mapped features. Here you point out a reduction in features.
We believe the text already addresses this apparent contradiction. The SfM dataset resulted in the most mapped tectonic landform indicators, but the participant used these to refine rather than expand their fault trace interpretation, retaining the primary fault and removing most secondary faults.

416-417: do you have a sense of if this revision was caused by "there must be something there-someone else saw it", vs "ahh OK, I can see it now"?
Participant reports indicated that the dominant effect of consulting the geologic map was validation of existing interpretations, which is already described in the preceding sentences.

417: earlier you used a hyphenated version of this. Also, do you mean "approach" or "method" or "procedure"?
L477: We changed "setup" to "experiment" for consistency with terminology used elsewhere in the manuscript.

424: do you mean the impact of all these things on epistemic uncertainty?
L483-4: We clarified point (4) to specify that our findings carry implications for the quantification of epistemic uncertainty in fault mapping, making it clearer that this is a distinct implication rather than an umbrella term for the other three points.

431a: average? Or the range of? I think this is consistent with the range up to n meters.
L496: We changed "average" to "reported" to more accurately reflect that the 50 m buffer falls within the range of fault-to-rupture separations reported across multiple empirical studies in Table 1, rather than representing a single average value.

431b: shoudl be table 1
L496: Thank you for catching this typo. We deleted one 1.

433-434: redundant
L498-9: We agree this phrasing was redundant. We simplified the sentence to "These results underscore how the interpretation of mapping performance is scale-dependent across different fault mapping applications."

435: for
L500: We made this change.

437: perceived?
L502: We added “perceived” uncertainty.

441: this could use a bit of elaboration
L506-8: We elaborated on why engaging with the entire map extent improves mapping, clarifying that regional context helps mappers correctly interpret local features and avoid missing faults outside areas of obvious expression. This is supported by our Fukushima-Hamadori results, where the participant who looked beyond the mapping area to identify the regional trend of the ranges produced the most accurate map.

448-450: this woudl be expected, I think. But it also presumes experienced mappers did not make (any) errors. I woudl say something along the lines of "undergradautes make more errors than experienced mappers" or even better, quantify that. Also, I think defining "error" somewhere is helpful. For example, if a mapper maps a lineament that is clearly tectonic but the next rupture does not occur on that same trace, is that an "error"?
L522-5: We agree that "errors" needed to be defined and that the original sentence implied only undergraduates made mistakes. We revised the sentence to clarify that "errors" refers specifically to misinterpretation of non-tectonic features as fault traces, such as fluvial scarps, and that while undergraduates did this more frequently, mappers at all experience levels made some misinterpretations.

453: I don't understand what this means
L528: We agree that "expected location quality and data density" was unclear. We replaced it with "expected level of detail and precision" to more clearly convey that mappers trained on fixed-scale maps have internalized what level of mapping detail is appropriate for a given scale.

454: I woudl replace this with digital since many air photos are very high resolution but still on paper.
L529: We made this change.

455: this can also mean zooming in and out too much. I think what you mean to say is that there is an optimal scale of observation for different landforms.
L529-31: We agree that the original phrasing was unclear. We revised the sentence to convey that each landform type has an optimal scale of observation, and that excessive zooming can be counterproductive, consistent with the suggestion.

457-463: this discussion of Bias is important, but I think it needs to move up in the manuscript, to a prior section. And then this sort of language can be used throughout the doc.
We believe the cognitive bias discussion is most appropriately placed in the discussion section, where it can be interpreted in the context of our results. Moving it earlier would introduce concepts before the reader has seen the data that motivate them. We did, however, revise "similar projects" to "similar fault systems or settings" to be more specific (Reviewer 1's suggestion) on L537.

466: often
L542: We made this change.

469: qualify this. Some experience is necessary. I would say all the undergards have some level .
L545: We agree that the statement needed qualification. We changed "experience" to "extensive experience" to acknowledge that while all participants had some level of geologic training, additional experience beyond a basic level did not improve mapping in settings with a single dominant landform.

479-480: this is a tough one to stomach. First off, I woudl suggest qualifying this. For example, "Data from our experiment do not support the hypothesis that experience improves interpretation reliability in areas of subtle tectonic geomorphic signals".
L555-7: We agree that the original statement was too absolute. We revised it to more precisely reflect what our data support, noting that our results do not support the hypothesis that experience improves fault mapping in areas with weak geomorphic signals or extensive anthropogenic activity, rather than making a universal claim.

480: rather than requiring the readers to refer to supplemental info for something that could be explained in a few words. I assume you are referring to local landscape height about active channels?
We believe this is already explained in the text — the sentence notes that tectonic landforms are subtle because topographic relief is only 24 m, which we believe sufficiently describes the landscape character without requiring the reader to consult the supplement.

485: may also be worth noting it is not just new ruptures that fall into the aleatory realm, but also faults that ruptured in the past that don;t rupture in each subsequent event.

L563-5: We agree that aleatory variability encompasses not only new ruptures but also previously active faults that do not rupture in each subsequent event. We revised the sentence to include both cases and corrected "aleatoric" to "aleatory" for consistency.

487: that had not ruptured in the 1892 earthquake.
L566-7: We made this change.

488: in the intro to this section you said it does improve it. It is a big jump to go from there to saying it does not ensure it. I think you need some discussion about where it can, and where it does not improve fault mapping
L568-9: We added a clause acknowledging that experience can improve mapping in some settings, bridging the transition from the earlier examples where experience did help to the cases discussed here where it did not ensure accuracy. We believe the existing subsections "What are the limitations of experience?" and "How and when does experience enhance mapping accuracy?" already provide the thorough discussion the reviewer requests.

489: also geomorphic, etc. and other subdisciplines. I suggest generaling this to avoid having to list all the sub disciplines.
L570: We agree that "structural knowledge" was too narrow. We generalized this to "geologic knowledge" to encompass the range of relevant subdisciplines.

490: THis seems like a bit of a jump. I am not sure your data support this interpretation. If the data do support it, I think it is worth a paragraph or two discussing it. I think that woiudl be a very important result if it is supported.
L570: We agree that over-reliance on heuristics is a general observation from the literature rather than a direct conclusion from our data. We hedged the statement by changing "remain susceptible" to "may be susceptible" to reflect that this is drawn from prior work rather than a novel finding of our study.

490-491: revise this sentence.
L572-3: We revised this sentence for clarity.

498: in these two settings, trench nearly perpendicular...
L585: We made this change.

500: Just an opinion but I think each heading deserves text.
We appreciate the reviewer's suggestion. We chose to retain the current structure as stand-alone subheadings are used consistently throughout the manuscript.

504: dense veg also obscures SFM (and lidar to a milder extent). I think you should clarify what is meant here.
L593-4: We agree that dense vegetation obscures landforms across all dataset types, not only

optical imagery. We clarified the sentence to note that lidar and SfM are also affected, as demonstrated by the Fukushima-Hamadori case.

507: was it really this low of resolution?
Yes, the El Mayor-Cucapah lidar dataset is indeed 6 m/pixel. This is the 2006 INEGI Sierra Cucupah dataset (OpenTopography, 2012).

510: What site was this used on? Also Borah ruture?
L600: We clarified that the SfM data referenced here was from the Southern San Andreas fault mapping area, which had the highest resolution data in our study. We added this specification in parentheses to avoid confusion with the Borah Peak site.

514: sort of a broad statement for all thse results: I think these statement like "significantly" really would benefit from quantification, and adding additional context as well.
L604: We do not have statistical data to quantify the degree of change in fault trace locations after consulting the geologic map. We therefore replaced "significantly" with "substantially" to convey the same meaning without implying a formal statistical comparison. The results in Section 3.4 show that most fault locations remained unchanged and that participants made only minor adjustments at a small number of specific locations.

517: worth mentioning somewhere (ideally early on) is that many faults on geological maps are inferred, or drawn based on linements. Particularly Quaternary faults. So this may be a bit circular, since existing maps are essentially expert-derived
L609-10: We added a brief acknowledgment that some geologic fault maps are themselves derived from remote sensing and geomorphic interpretation, which can reduce their independence as a validation source for desktop mapping. “We note, however, that some geologic fault maps are themselves derived from remote sensing, which reduces their independence as a validation source for desktop mapping.” We framed this as a nuance rather than a limitation, consistent with our overall conclusion that geologic maps remain a valuable foundational resource.

521: geomorphic and desktop? Is this referring to the same method in this paper?
L612-3: We clarified this phrase to make explicit that "geomorphic" refers to the landform-based mapping approach and "desktop" refers to the remote, non-field-based method of conducting that mapping. These are two distinct aspects of our methodology and not redundant terms. We revised the sentence to "Geomorphic fault mapping conducted remotely has limitations" to avoid confusion.

528: in plenty of cases, geophysics results in significant cost savings. i woudl suggest revising this to say "if resources are available" rather than specifyig project budgets.
L619: We agree that framing this as a budget constraint was too narrow, as geophysical

investigations can in some cases result in cost savings. We replaced "if the project budget allows for these types of investigations" with "if resources are available."

531: do your data support this? If so, I would add that info, and what you are basing it on. "novice mappers reported that they appreciated the structured approach, whereas experienced mappers applied a more flexible approach, presumably due to their comfort with mapping".
L623-4: We agree that this claim should be supported by evidence. We added references to Figures 4, 5, and 6, which show that some less experienced participants produced maps of comparable quality to more experienced mappers across the three main repeat mapping areas. Detailed participant feedback on the structured approach is reported in Adam et al. (2025).

540: s
L631: We made this change.

554: I suggest discussing some of the caveats here, e.g. very small sample size, tectonic settings, etc.;
556: this may ruffle some feathers. Ultimately, experience will be needed to review and approve etc.
L647-51: We agreed with both suggestions. We added a caveat noting the small sample size and limited tectonic settings, and reframed the sentence about hiring experienced geologists to avoid unnecessarily blunt language. We combined both points into a single sentence for concision.

558: Delete “as more experience is not harmful”
L653: We made this change.

560-563: since you are getting into the commercial aspects of this stuff, I suggest modifying the section title to be more specific. Something like "considerations for selecting professional mappers"
L644: We made this change.

564: approaches
L660: We made this change.

564-565: influence differences in interpretation
L661: We made this change.

568: alternatively, co-mapping…
L664: We made this change.

569a: novices' since I think you are showing their possessing self assurance.; 569b: perhaps "reluctant" may work better than afraid, and I'd also suggest replacing the last part with "due to lack of confidence".; 569c: unnecessary and speculative unless you want to elaborate with an example
L666: We made three changes to this part of the sentence. For 569a, we added an apostrophe to "novices'" to correctly indicate possession. For 569b, we replaced "afraid" with "reluctant" as a more precise and professional term. For 569c, we replaced the speculative clause "possibly better, ideas" with "ideas due to lack of confidence" to ground the statement in a more defensible rationale rather than implying novice ideas are necessarily better.

570: Clarify what you mean here. Do you mean "correctly identify more fault traces"?
L667-8: We clarified "reveal more fault traces" to "improve the accuracy of fault trace identification" to more precisely convey that combining maps from multiple individuals leads to more reliable fault interpretations than any single mapper alone.

576: I'm not sure I understand what this part means. Are you referring to the uncertainty in the rupture zone? Or just an improved understanding of uncertainty?
L672-4: We agree the original phrasing was unclear. We revised the sentence to clarify that identifying settings with large epistemic uncertainty informs decisions about fault zone buffer widths in hazard assessments.

578: I think this needs more explanation. The widest mapped defomration zone? Of the zone with the largest variation of mapped traces?
L676-8: We agree the phrase "widest mapped zone" was ambiguous. We clarified that this refers to the widest spread of mapped fault traces across mappers, i.e., the zone of variation in fault trace placement, rather than the widest deformation zone.

581: I think there is some mixing of uncertainty with aleatory variabiity. The mapped trace of past fault ruptures may be acurate, but in some cases the latest rupture is not follwoing the mapped trace. That randomness is best modeled with aleatory variability.
L686-94: We address the mixing of epistemic uncertainty and aleatory variability in the paragraph added to Section 4.5 in response to Reviewer 2's fundamental comments, which clarifies that our reported values represent an upper bound on epistemic uncertainty as they incorporate some aleatory variability in rupture location.

588: I would say that quantifyication of the epistemic uncertainty is an important component of fault mapping, rather than sayig it provides a practical framework. I am not sure what the practical framework would be.
L695-7: We agree that "practical framework" was vague. We revised the sentence to state more directly that quantifying epistemic uncertainty is an important component of fault

displacement hazard assessments with direct engineering applications, consistent with the reviewer's suggestion.

590: Delete “(ideally)”
L701-2: We made this change.

591: I agree, but how would this happen? What methods would you suggest for dealing with the aleatory component?
L702: We agree that the aleatory component needed more context. We noted that aleatory variability in rupture location can be characterized through probabilistic fault displacement hazard analysis (PFDHA), citing Youngs et al. (2003), which provides the methodological framework for doing so. A full treatment of PFDHA methods is beyond the scope of this paper but we believe this reference gives the reader a path forward.

592: the mean and median are not really zones, but lines, right? Are you trying to state that their epistemic uncertainty zones may be narrower due to greater precision (lower sigma)?
L704-705: We agree that the mean and median are single values rather than zones. We revised the sentence to clarify that sites with well-defined scarps may justify using a narrower uncertainty buffer based on the mean or median epistemic spread.

600: again, looking back, were these ruptures first of thier kind, or where the missed landforms/ I think in the Thompson-jobe paper there was some discussion that landforms were not only missed, but may not have been detectable in the first place.
L712-4: We agree that some missed rupture locations may reflect landforms that were not detectable in pre-event remote sensing data or ruptures with no prior surface expression, rather than mapping errors. We added a clause acknowledging this caveat with reference to Thompson Jobe et al. (2020), which discussed this issue in the context of the Ridgecrest earthquake sequence.

603: While I like this result, there are a bunch of places throughout the text that say otherwise.
In context, this statement is complementary rather than contradictory to our earlier findings. The preceding sentences establish that systematic mapping can help less experienced mappers produce quality results, while this sentence acknowledges that map quality is still influenced by expertise, landscape character, and data type and resolution. These are complementary rather than contradictory points. Systematic mapping reduces but does not eliminate the influence of experience on map quality.

606: I think "quantifying" is more appropriate than reducing. Many instances can lead to larger uncertainty estimates, and the goal of any sort of seismic source model input is to quantify it. I think it is commonly written as "reduce uncertainty" but that is, in my experience, not always

the result or even goal.
L720: We made this change.

Figure Comments

Figure 2 caption: i think earlier you mentioned you adopted the Saramiento language of principal and distributed
We use "primary and secondary" to describe mapped fault traces, consistent with the confidence ranking system described in the mapping approach, and "principal and distributed" to describe coseismic ruptures, consistent with Sarmiento et al. (2021). Figure 2 correctly uses "primary or secondary ranking" as it depicts mapped fault traces rather than coseismic ruptures. We believe this distinction is already clear in the manuscript.

Figure 7: when I look at the inset on Fig 7C I am reminded of how fine-scale this mapping. I am wondering how much of the difference is a function of how careful a mapper is vs incorrectly interpreting rupture locations.
We agree this is an interesting observation. The differences in mapped fault locations at this fine scale reflect both interpretive variability (epistemic uncertainty) and mapper carefulness, which are difficult to separate. This is discussed in Section 4.5 where we quantify the spread of mapped traces as an upper bound on epistemic uncertainty.

Figure 10D: this may be an example of what I was referring to. I am not too sure, but were there any idications that this highly arcuate section could have been mapped/predicted correctly? Or we this, and other similar complexities simply surface effects of the rupture, things that coudl not have really been accurately predicted?
The arcuate section at Borah Peak likely reflects surface effects of the rupture rather than a mappable pre-existing feature, consistent with our discussion in Section 4.3.2 where we note that ruptures outside the 2σ buffer may reflect aleatory variability in rupture location rather than mapping error.

Figure 10 caption: ultimately I am getting the sense that this is realy about accurace and precision. your 2 sig epistemic zone is more a reflection of the precision of mappers, and the accruacy piece can come from either proximiltiy to the paleo-scarps, or proximity to the MRE scarps.
We appreciate this framing of the results in terms of accuracy and precision. The conceptual distinction between precision (spread of mapped traces, reflected in the epistemic uncertainty zone) and accuracy (proximity to the actual rupture location) is addressed in the discussion, particularly in Sections 4.3 and 4.5, where we interpret our epistemic uncertainty values in the context of mapping performance and hazard applications.

Dear Randy Williams and Reviewers,

Thank you again for your handling of the review process for our manuscript, "The impact of mapper experience and data on the quality of geomorphic fault mapping." We are grateful for your second look. This second round has further improved the manuscript.

We have addressed Reviewer 1's comments below in-line in purple. We did not include line numbers in most of our responses because the lines do not differ by more than one line from the reviewer's comment.

Our most substantial change to the manuscript was to address the time since the most recent surface-rupturing earthquake at each study site. We added the MRE dates and a small addition to the discussion. This was an important point, and we appreciate Reviewer 1 for bringing it to our attention.

Sincerely,
Malinda G. Zuckerman and coauthors

Dear Malinda G. Zuckerman, Chelsea Scott, Ramon Arrowsmith, Rachel Adam, Chris Madugo, Rich D. Koehler, Albert Kottke, Hans AbramsonWard, Eldon Gath, Tania González, Brian Gray, Ozgur Kozaci, Thomas K. Rockwell, Trevor J. Austin, Tamarah King, Rachel Micander, Ethan Leuchter, Ian Pierce, Erin Alexander, Aube Gourdeau, Maximilien Laly, Izabelle Ogilvie, Sophie D. Rothman, Dominik R. Vlaha, Elaine K. Young, Veronica Prush, Christie Rowe, Madeline Schwarz:

I hope this email finds you well. I have reached a decision regarding your submission to Seismica, "The Impact of Mapper Experience and Data on the Quality of Geomorphic Fault Mapping ". Thank you once again for submitting your work to Seismica.
Based on reviews I have received, your manuscript may be suitable for publication after some minor revisions.

Both reviewers are generally satisfied with the revisions and your responses to their comments. As you will see, however, Reviewer #1 feels that a few small, high-value additions would make for a more complete contribution. Please address these comments in a revised version.

I do not anticipate sending this manuscript back out for peer review. In addition to the items listed below, please also upload publication ready originals of all figure files once you have addressed the final comments.

When you are ready to resubmit the revised version of your manuscript, please upload:

- A 'cleaned' version of the revised manuscript, without any markup/changes highlighted.
- A pdf version of the revised manuscript clearly highlighting changes/markup/edits.
- A 'response-to-reviewers' letter that shows your response to each of the reviewers' points, together with a summary of the resulting changes made to the manuscript.

Once I have read your revised manuscript and rebuttal, I will then decide whether the manuscript either needs to be sent to reviewers again, requires further minor changes, or can be accepted.

If you deem it appropriate, please check that the revised version of your manuscript recognises the work of the reviewers in the Acknowledgements section.
Please note that Seismica does not have any strict deadlines for submitting revisions, but naturally, it is likely to be in your best interest to submit these fairly promptly, and please let me know of any expected delays.

I wish you the best with working on the revisions. Please don't hesitate to contact me with any questions or comments about your submission, or if you have any feedback about your experience with Seismica.

Kind regards,
Randy Williams

Reviewer 1:
This is a re-review of “The impact of mapper experience and data on the quality of geomorphic fault mapping” by Zuckerman and coauthors. The authors have considered the reviewer comments, and carefully revised their manuscript. The reorganization of the manuscript and the improved definitions and separation of epistemic and aleatory uncertainties are especially appreciated, and I think the message is much clearer. The results and interpretations will be of interest to a broad audience focused on active fault mapping and seismic hazard. I have one remaining broader comment, expanded on below, followed by a few minor comments by line number for clarification. This work represents an important step forward toward understanding uncertainty in fault mapping and its implications for seismic hazard, and I look forward to seeing this work published.

Major comment:

1. I appreciate that the authors added a mention of the influence of the timing of the most recent earthquake on the landscape geomorphology, but I still think it would help the reader to better understand some of the context and discrepancies in the mapping if it was discussed specifically for each site. For example, a fault that had surface ruptures only several hundred to a few thousand years ago may have better landscape expression than a fault with an MRE of a few thousand to ten thousand years, or alternatively, faults with shorter recurrence intervals may have better landscape expression than faults with longer recurrence intervals. This will directly translate to how well participants of any experience level will map fault traces, and is worth mentioning, if only for a few sentences when providing background information on the faults, to provide the readers some context for interpreting the results.
   We also appreciate this fine point. Please see our response to the comment at L149.
   In addition to those changes, we also added a couple of sentences to the discussion in the subsection, *Overview,* where we state that participants did indeed have an easier time mapping and predicting ruptures in the area with the

most recent MRE and a more difficult time in areas with a more distant MRE (L472). This is an interesting result, and we are glad for the opportunity to include it.

Minor Comments:
L115, L119: Some text strike-throughs and additions – please check that the correct Sections are referenced here.
Yes, thank you for catching this typo. The changes are correct; they simply weren't cleaned up.

L145: Define Mw at its first appearance in the manuscript.
We made this change.

L145-147: I suggest adding the country to the earthquakes in this list for those unfamiliar, and referencing Figure 1.
We made this change.

L149: Consider adding the time since the most recent surface-rupturing earthquake here for readers unfamiliar with the details of each earthquake. Or the authors could include it in the results of each mapping area.
We added these dates in the results:

Borah Peak on L265 (5-7 ka)
El Mayor Cucapah: added more specific language on L297 (118 yrs)
Fukushima-Hamadori: L322 (800-6200 yrs)
Ridgecrest: L381 (5-7 ka)
SAF: already included on L402 (300 yrs)

We also added a clause on L150 that includes the range of years, emphasizing the differences in the time since the most recent surface-rupturing earthquake.

L268: it may be useful to state this in map distance instead of just elevation difference, as the previous section was focused on the fault zone widths using horizontal distances.
Yes, good point. We made this change and led with the map distance.

L366: Aren't triangular facets typically also associated with normal faults?
Yes. We added the clause ", despite triangular facets being more commonly associated with normal faulting" at the end of the sentence. This addresses your point that the participants mapped the triangular facet and still classified the fault as strike-slip even though those are typically evidence of normal faulting.

L406: Define NAPP at its first appearance in the manuscript.
Good catch. I see I forgot to define it in the main text; it's only in the caption. We made this change.

L483: Or mapping features that are apparent fault scarps and not actually faults and will not rupture.
We added in parentheses, "or apparent faults," to account for those features that are not tectonic.

Recommendation: Revisions Required

Reviewer 2:
I appreciate the opportunity to have reviewed "The impact of mapper experience and data on the quality of geomorphic fault mapping" by Zuckerman et al. The study has clear implications for seismic hazard studies, and it is clear that the authors have thought critically about the role of experience while also evaluating how different data resolutions and geologic maps impact mapping quality.

I believe the authors did a thorough and appropriate job of addressing the two broad comments I had made in my initial review. They also thoughtfully addressed each of my specific comments, and the overall reorganization of the manuscript structure is quite helpful and improves the clarity of the text.

By utilizing a repeatable desktop experiment with mappers ranging from undergraduate students to senior professionals, this paper provides important data on the limits of experience and the nature of epistemic uncertainty. It is also useful that the authors describe how mappers uniformly struggle in specific types of terrain and identify the most useful forms of data that support high quality mapping.

This work represents a valuable contribution to applied fault mapping for seismic hazard studies. I expect that this work will support the improvement of, and critical understandings of uncertainties development of seismic source models, particularly for probabilistic fault displacement hazard analyses.
Recommendation: Accept Submission