Response to Anonymous Referee #2 Comments:

We thank the referee for their valuable comments, and here are our point-by-point responses. Our responses are in red, and note that RC = referee comments, and AC = author comments. However, before the point-by-point response for each reviewer, we would like to provide a General Response, which is identical for both reviewers, that should help to address some of the overarching and overlapping concerns that were brought up by both Reviewers 1 and 2.

AC: General Response

In reading through the comments from both reviewers, we saw that there was one main issue that weaved throughout, namely the need for a more thorough discussion of the study's contribution to drought feedback loop, along with the associated limitations and caveats. To address this, we added a more detailed discussion in the Conclusions section, as well as a new figure to the manuscript, now "Figure 9", below.

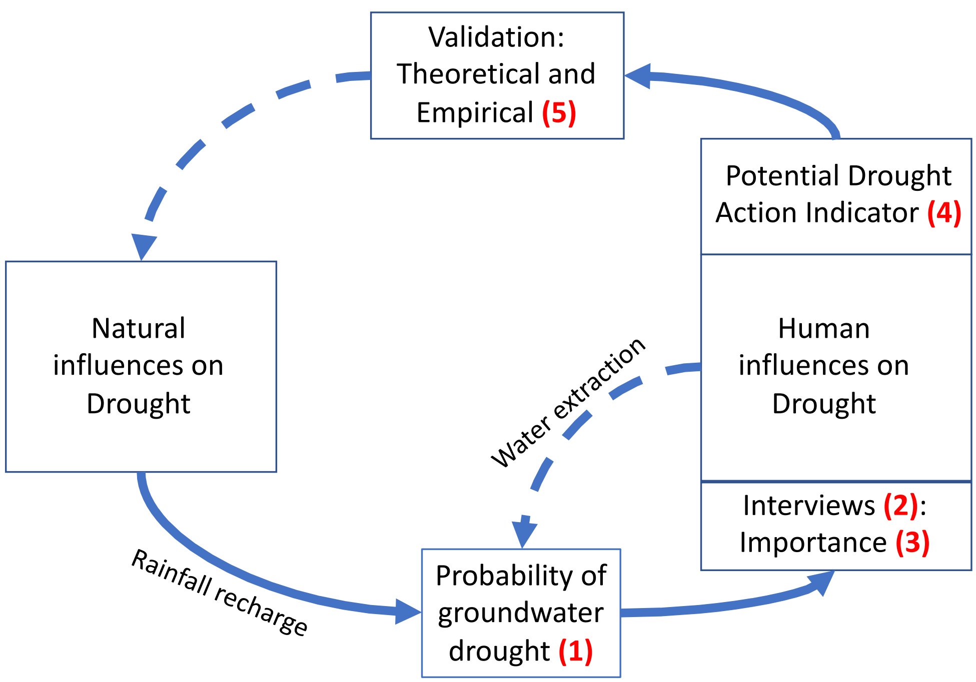

Figure 9. Conceptual map of drought feedback loop components addressed in this study (blue solid lines) and remaining gaps (blue dashed lines). Numbers correspond to discussion points in Conclusions.

Here we provide the excerpt that will be added to the Conclusions section of the manuscript to describe the new Figure 9 and the study limitations:

"We develop and demonstrate this methodology as a step towards closing the drought feedback loop, but note that there are caveats and limitations that warrant discussion. A conceptual overview of the contribution of our study to the drought feedback loop is shown in Figure 9, and we use this figure to identify five places where there is scope for future enhancements; each number below corresponds to a place in the drought feedback loop in Figure 9:

1) For the natural influence on drought, we examine the probability of groundwater drought. In our case, the groundwater levels are closely related to rainfall recharge, which is a function of natural climate variability. We recognize that this is not the case for many groundwater aquifers, where human activities, such as groundwater extraction, may trump the natural climate signal (e.g., Tarhule and Bergey 2006), often leading to water scarcity, rather than a natural phenomenon of temporary water deficiency. In many systems a full water balance would need to be examined to understand the relative contributions of extraction versus moisture deficit to the likelihood of going below a relevant hydrologic threshold.
2) Our interviews were conducted following a drought event, and we recognize that the timing of the interviews will likely affect the responses. For instance, interviews conducted during wet or average conditions might elicit less polarized responses, since drought impacts haven't been recently experienced. We note that our approach of applying the interview responses across different climate conditions (i.e., wettest to driest) makes the assumption that the importance of water uses and management preferences are stationary. We acknowledge that different climate conditions, as well as cultural change, technological innovation, climate adaptation, and other processes are likely to influence the cultural factors we investigated here and may mediate how people interact with their environments. Future work could investigate how responses change with different climate conditions over time, and the subsequent implications for drought action. However, hazards and disasters research is almost always conducted immediately after an event, so this is a wise-spread epistemological issue with both pros and cons in terms of what we learn from post-disaster research.
3) We use stakeholders' importance ratings as a proxy for their willingness to take action in relation to particular water uses, where by "action", we generally mean some effort towards drought mitigation. The interviews included questions about the importance of different water uses to test the application of the Cultural Theory of Risk (usually applied in a more global sense) to a specific water management issue, which had not been done before (Lazrus 2016). For the purposes in this paper, multiplying the importance ratings by the probability served as a way to make an objective characterization of drought subjective, that is, we wanted to modulate the groundwater drought probability by each individual stakeholder's lens.
4) The formulation of the PDAI strongly affects the conclusions drawn. Our formulation of the PDAI follows from other precedents in risk management that take the product of the likelihood of an event and its importance (Jones and Preston 2011; Oppenheimer et al. 2014). However, the functional form of the PDAI is flexible, allowing it to be tailored to other locations. As such, we note that the PDAI, as well as the best data to use to calculate it, will depend on the needs of the community, as well as the water system context.
5) We use social science theory to interpret our results, and to better understand the theoretical underpinnings of how and why people take action in response to drought. However, we note that empirical validation is important for indicator development and refinement. We recommend that future project designs include a validation component in the methodology. This could take the form of follow-up interviews, such as direct

feedback from stakeholders on if the indicator reflects their willingness to take action for certain water uses at certain drought levels. Methods, including stakeholder processes, for developing and evaluating drought indicator effectiveness have been put forth in the drought community (Steinemann and Hayes 2005 Steinemann and Cavalcanti 2006, Steinemann et al. 2015). Other options for validation can be indirect, such as looking at historical data, like government and local reports, media, and/or other collected response information, e.g., in the United States the US Drought Impact Reporter (Wilhite et al. 2007). Promising methods for mining social media, such as Twitter, have also been developed (Demuth et al. 2018) and could be adapted for evaluative purposes.

Although the methodology to develop the PDAI is experimental, we posit that explicit efforts to combine natural and human perspectives is critical to gaining a deeper and more nuanced understanding of drought feedbacks, and this paper provides a novel contribution to this end."

We will refer to this General Response in our point-by-point responses, below.

Interactive comment on "Characterizing the Potential for Drought Action from Combined Hydrological and Societal Perspectives" by Erin Towler et al.

Referee 2.

RC0. The manuscript by Towler et al. investigates the potential for drought action. The stated "goal of this paper is to provide an experimental methodology towards a better characterization of several components of the drought feedback loop" and the study claims to have done this by "developing an index to characterize how natural influences on drought inform potential human actions on drought." The general topic is important and deserves innovation and systematic research. Unfortunately, I do not really see this was achieved by the presented material. I found the hydrological analysis particularly weak and the relation between drought events, general wetness/dryness, and potentially water scarcity rather than drought defined and elaborated too little for a hydrology journal. The manuscript has some technical issues and lacks a thorough discussion section on uncertainties, biases, and comparing outcomes with other studies. Nevertheless, the material is interesting and an improved version could make a valuable contribution to the topic. My overall assessment is that the manuscript may be more suitable to a journal with less demand on the hydrological science part and perhaps more focus on water resources management or hazards. I hesitate to recommend its publication in HESS.

AC0. We appreciate the reviewer's comments. We address some of these overarching concerns in our General Response above, as well as the more specific comments in our point-by-point responses below. Here, we would like to respond to the reviewer's comment on journal suitability. To this end, we provide an excerpt from the aims and scope of HESS, using bold to highlight the parts of the scope that appealed to us in our selection of this journal for this manuscript: "HESS encourages and supports fundamental and applied research that **advances the understanding of hydrological systems**, **their role in providing water for ecosystems and society**, and the role of the water cycle in the functioning of the Earth system. **A multi-disciplinary approach is encouraged that broadens the hydrological perspective** and the advancement of hydrological science through integration with other cognate sciences and **cross-fertilization across disciplinary boundaries**. HESS, therefore, **aims to serve not only the hydrological science community but all earth and life scientists, water engineers, and water managers**, who wish to publish original findings on the **interactions and feedbacks between**

the governing processes of the water cycle….". The selection of a journal is subjective, but from the excerpt above, we believe that our paper lies squarely within the bounds of HESS.

Major comments

RC1) The scientific frame needs to be laid out more specifically. What exactly is meant by Drought Action (title). With the stated goal repeated above, I would first expect a clear definition of and thorough elaboration on what is called the 'drought feedback loop' in the stated goal. To me it remained unclear what is meant by that as well and what it has to do with this study - or where and why exactly is the research gap that is addressed here. In this context, the introduction and discussion ignore literature and existing experience on drought risk management and drought plans and quantitative trigger levels developed with stakeholder processes elsewhere.

AC1. We acknowledge that there are different ways to approach the scientific framing of a drought problem, and have taken steps to better specify how our work fits within our selected frame. As noted in the General Response, above, we saw an opportunity to address this directly in the Conclusions, where we now do a more comprehensive job of identifying our contribution in the drought feedback loop, as well as acknowledging the study limitations; this occurs in Figure 9 and the five corresponding points in the Conclusions (see General Response). Further, we have elaborated on our goal in the Introduction (additions in bold):

"The goal of this paper is to provide an experimental methodology towards a better characterization of several components of the drought feedback loop, **specifically to gain understanding on how and why people might take action in response to drought**."

Another point raised here speaks to the vast drought literature that exists, and the need to better connect with existing experience on drought risk management, drought plans, and quantitative trigger levels developed with stakeholder processes elsewhere. We have added two references and changed a sentence in the Introduction to better illustrate this point; if the reviewer has other specific reference suggestions, please let us know:

"The need for more proactive drought planning has led to increased interest in the development of drought management plans (e.g., **Wilhite et al. 2000**, Wilhite et al. 2005, Knutson et al. 1998). **Drought risk management requires identifying drought indicators and triggers (Steinemann and Hayes 2005), which can be developed and evaluated using stakeholder processes to make them useful for decision-making (Steinemann and Cavalcanti 2006, Steinemann et al. 2015)**."

However, we also note that our approach is distinctive from stakeholder processes that are explicitly designed to develop a drought plans and quantitative triggers. Rather, one unique aspect of our contribution is to use the social science theory to understand the theoretical underpinnings of how and why stakeholders take action in response to drought. We add text to this effect here (addition in bold):

"Though we do not directly validate the PDAI, we are able to interpret the findings to provide insights to water management policy using additional interview data on stakeholder worldviews and social science theory; **this is unique in that it allows us to investigate the theoretical underpinnings that are not typically explored in drought risk studies**."

We also add this in point 5 of the Conclusions (also in "5" in "General Response" above):

"5. We use social science theory to interpret our results, and to better understand the theoretical underpinnings of how and why people take action in response to drought. However, we note that empirical validation is important for indicator development and refinement. We recommend that future project designs include a validation component in the methodology. This could take the form of follow-up interviews, such as direct feedback from stakeholders on if the indicator reflects their willingness to take action for certain water uses at certain drought levels. Methods, including stakeholder processes, for developing and evaluating drought indicator effectiveness have been put forth in the drought community (Steinemann and Hayes 2005 Steinemann and Cavalcanti 2006, Steinemann et al. 2015). Other options for validation can be indirect, such as looking at historical data, like government and local reports, media, and/or other collected response information, e.g., in the United States the US Drought Impact Reporter (Wilhite et al. 2007). Promising methods for mining social media, such as Twitter, have also been developed (Demuth et al. 2018) and could be adapted for evaluative purposes."

Wilhite, DA, Hayes, MJ, Knutson, C, Smith KH, Planning for drought: Moving from crisis to risk management. Journal of the American Water Resources Association, 36(4), p. 697-710. DOI: 10.1111/j.1752-1688.2000.tb04299.x, 2000.

Steinemann, A. C, Iacobellis, S. F., and Cayan, D. R.: Developing and Evaluating Drought Indicators for Decision-Making. *Journal of Hydrometeorology*, *16*(4), 1793–1803. http://doi.org/10.1175/JHM-D-14-0234.1, 2015.

RC2) Overall, the material is presented very much from a descriptive case study perspective, starting with a long description in Background. Each subsection in 3. also starts with a narrative of the case study region's conflicts etc., rather than theoretically presenting the approach and then briefly stating the data of the case study used to illustrate the approach. An international readership as in HESS will be interested in this, not in the case study details.
AC2. The goal of the paper is to put forth a methodology that is developed and demonstrated for a particular case study. Every application will have unique details, but we believe that the details matter. For example, we note in the paper work by Sivapalan (2012), who suggests that one needs to study a particular water system in detail to understand human-hydrology interactions. However, the reviewer makes a good point that we need to explicitly address how this approach can be tailored to other locations. We address this in point 4 of our Conclusions (also listed in General Response Point 4 above):

> "4. The formulation of the PDAI strongly affects the conclusions drawn. Our formulation of the PDAI follows from other precedents in risk management that take the product of the likelihood of an event and its importance (Jones and Preston 2011; Oppenheimer et al. 2014). However, the functional form of the PDAI is flexible, allowing it to be tailored to other locations. As such, we note that the PDAI, as well as the best data to use to calculate it, will depend on the needs of the community, as well as the water system context."

RC3) The manuscript repeatedly states that the study takes a hydrological view on drought. Perhaps my most substantial criticism is that the reader does not receive this hydrological view. As mentioned in the 'Background', the case-aquifer is rain-recharged and feeds springs and rivers as well as groundwater extraction - hence the aquifer's water balance is crucial. A hydrological perspective would need to provide rain and recharge data (or at least climatic water deficit) time series, spring and river flow data as support for when there is drought, and an assessment how the groundwater levels are affected by abstraction as compared to the natural signal (if the van Loon et al. perspective is taken, proof is need what type of drought is considered here exactly). Are there trends - it looks like it? Together with some hydrogeological information, all this is missing and hence I do not see how the occurrence of drought (from a hydrological view a natural phenomenon of temporary water deficit that occurs rarely) can be distinguished from water scarcity or overexploitation. All this needs to be analysed in detail to know what it is exactly that one is feeding into such an index as the one created.

AC3. This is an important comment that has helped us to see how we need to clarify our "piece-of-the-puzzle" on the hydrological perspective. A key point we tried to get across was that the hydrology of this system has been extensively studied elsewhere, and we took the approach of referencing those studies and using the key conclusions in the development of the PDAI. We have added a more specific sentence on the previous study by Christenson et al. 2011 in the Background: "The study included a water balance, hydrogeological study, and groundwater model of the aquifer, and shows that although water is extracted, groundwater pumping from the aquifer is relatively small, and that the groundwater-fed streamflow discharge is mostly related to rainfall recharge (Christenson et al. 2011)." As such, this allows us to use the probability of groundwater drought as the natural influence of drought in our drought feedback loop, as is now seen in the new Figure 9. We point this out in our new Figure 9 and the limitation of this in Point 1 of our conclusions:

> "1. For the natural influence on drought, we examine the probability of groundwater drought. In our case, the groundwater levels are closely related to rainfall recharge, which is a function of natural climate variability. We recognize that this is not the case for many groundwater aquifers, where human activities, such as groundwater extraction, may trump the natural climate signal (e.g., Tarhule and Bergey 2006), often leading to water scarcity, rather than a natural phenomenon of temporary water deficiency. In many systems a full water balance would need to be examined to understand the relative contributions of extraction versus moisture deficit to the likelihood of going below a relevant hydrologic threshold."

Similarly, we reference some of our previous work – the Towler and Lazrus 2016 – that shows the empirical relationship between groundwater levels and a spring in a Recreation area. For the benefit of the reviewer, we include one previously published figure here from that paper that shows where the Recreation thresholds are derived from:

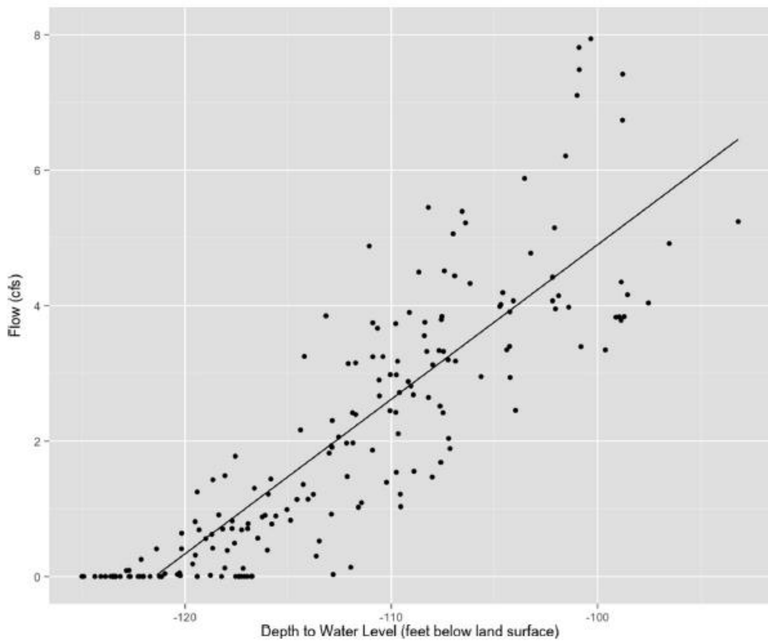

Fig. 2. Scatterplot of monthly Fittstown groundwater depths and Antelope Springs flows ($r = 0.87$).

We also shared this figure with the other reviewer, as they brought up a related point in asking about the -120 feet extreme/new normal threshold, the response to which we share here: The x-axis is the same groundwater level data used here, and this is related to spring flows in a Recreation area. From the graph, between -117 and -120 feet below the surface, there is variability in the springs flow, but the linear fit shows a monotonic decrease between the two levels. To clarify this in the paper, we say: "The second threshold is the "severe" threshold: this is when the groundwater level lowers further, to 117 feet below the surface, which is the level at which artesian springs in the area have **minimal flow or stop flowing altogether**, affecting uses such as wildlife and recreation. **For illustrative purposes, we also look at an "extreme" threshold of groundwater levels to 120 feet below the surface, which have been experienced in the aquifer and further the likelihood of minimal or stopped spring flows (see Figure 2 in Towler and Lazrus 2016**)."

RC4) With the stakeholder process published previously and the very limited hydrological analysis based on one groundwater well record only, the main argument of novelty is the PDAI. For me it was not clear how "potential drought action" is linked to the 'importance' interviews (theoretically). I did not have time to read the cited publication on the stakeholder interviews, but I think the infos given here are not sufficient to understand and follow the argument for a PDAI. Generally, I am not at all convinced about the introduction of yet another index on drought as there is enough confusion over existing drought indices already. Some justification why this is an index (and not just called what it is - function of...are there precedents in other hazards?) and a thorough assessment of transferability and usefulness beyond this case would be needed to justify this as the main contribution to the current debate on the topic.
AC4. We see your point about the proliferation of drought indices, and we want to be sure that we are clear on our study contribution and limitations. First, we have changed the "I" in PDAI from "Index" to "Indicator"; although subtle, we feel that an index is more of a direct measure of something, where in indicator reflects the a more indirect, but insightful, measurement. More

concretely, we have also added our conceptual figure, Figure 9, of the drought feedback loop a discussion in point 3 about the Importance ratings:

"3. We use stakeholders' importance ratings as a proxy for their willingness to take action in relation to particular water uses, where by "action", we generally mean some effort towards drought mitigation. The interviews included questions about the importance of different water uses to test the application of the Cultural Theory of Risk (usually applied in a more global sense) to a specific water management issue, which had not been done before (Lazrus 2016). For the purposes in this paper, multiplying the importance ratings by the probability served as a way to make an objective characterization of drought subjective, that is, we wanted to modulate the groundwater drought probability by each individual stakeholder's lens."

Specific or technical comments

RC5) Equations are not numbered and variables are not explained/defined consistently. Unnecessary use of multi-letter variable names (use z with various subs for gw levels and provide units, etc.). Please see HESS instructions for manuscript preparation regarding mathematical notation, use of equations, symbols, etc..

AC5. Thank you. We have fixed and labeled the two equations in the manuscript.

We express the PDAI as a function, *f*, of (i) the decadal probability (P) of the groundwater level, Z, exceeding the hydrologic threshold, z, and (ii) the importance ratings (I):

$$PDAI = f(P(Z<z), I) \qquad \text{(Equation 1)}$$

Here, we define *f* as the product (i.e., multiplication) of the two explanatory terms:

$$PDAI = P(Z<z) \times I \qquad \text{(Equation 2)}$$

RC6) L. 277 Why smoothing by a 10-year running window? Groundwater heads are already smoothed by the dampening processes in the hydrological cycle, but more importantly, any thresholds for management decisions and thus for the analysis will not use that, but actual water level. This requires justification in that respect.

AC6. We agree that we need to add information surrounding this selection. First, we want to clarify that our purpose was to look at the frequency (within a given time window) of a particular drought severity (here groundwater threshold exceedance). We have generalized the approach in the methods: "To quantify the threshold exceedance, we calculate the percent frequency of exceedance[1] for each threshold in the historical record. To calculate the exceedance frequency, a time window needs to be selected; we initially examined 5-, 10-, 15-, and 20- year windows. Specifically, we calculate the number of months during each x-year running window that the threshold was exceeded across the available record. For example, for the 10-year window, it would be 1959-1968, 1960-1969, etc., all the way to 2003-2012. Henceforth, we refer to this as the groundwater drought likelihood."

Next, in the results, we now reference the different time windows, show the results in the Supplemental Material, but select 10-years for our results to strike a balance between shorter time windows (e.g., 5-years) that show large variabilities and greater time windows (e.g., 15+-year windows) where all of the variability is washed out:

"Groundwater drought likelihood is calculated as the number of months within each 5-, 10-, 15-, and 20-year running window that the level went below a particular threshold. Drought

---

[1] We note that groundwater threshold levels are negative; so here we define "exceedance" as going below (more negative) than the threshold.

likelihoods for the selected time windows (5-, 10-, 15-, and 20-years) are shown in the Supplemental Material, Figure S2. Results for each time window follow similar patterns, though as expected, the shorter the time window, the greater variability in the likelihood. We selected the 10-year running window for calculating the PDAI as it strikes a balance between shorter time windows that have high variability and likelihood swings (e.g., 5-year windows) and longer time windows (e.g., 15-, and 20-years) where much of the variability gets washed out. As such, Figure 3 shows the decadal likelihood for the moderate and severe threshold."

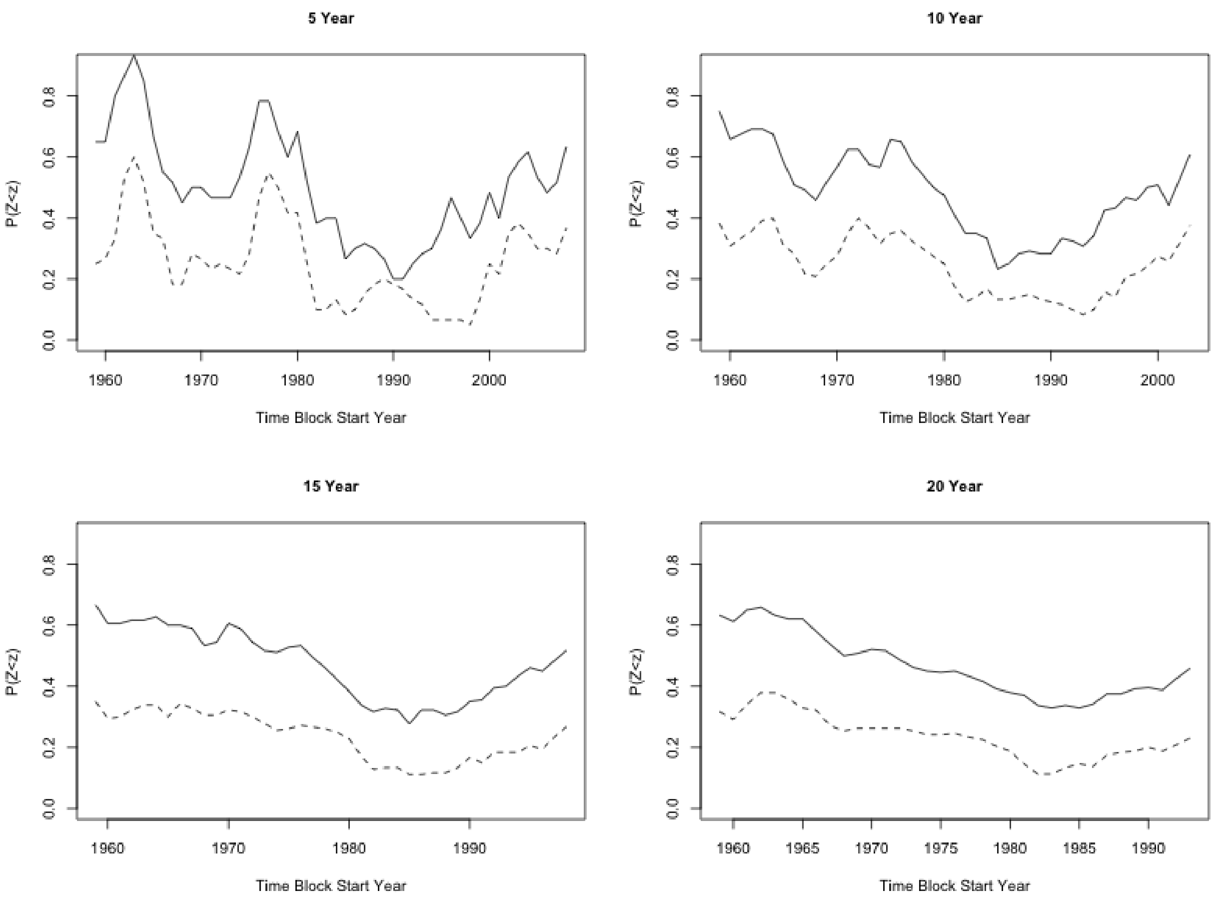

Figure S2. Groundwater drought likelihood (P) of the depth to groundwater level (Z) going below the moderate (z=-111 ft, solid line) and severe thresholds (z=-117 ft, dashed line) for selected time windows.

Also, as mentioned in AC1, we note that our approach is distinctive from stakeholder processes that are explicitly designed to develop a drought plans and quantitative triggers. Rather, one unique aspect of our contribution is to use the social science theory to understand the theoretical underpinnings of how and why stakeholders take action in response to drought. So here, we are not as tied to the actual water levels, and instead are looking at a frequency within a particular time window.

RC7) L. 281ff What exactly is r? Pearson correlation coefficient or some rank correlation? What other indices, etc.? All computations and data need to be introduced in the Data and Methods section. Not here.

AC7. Thank you for pointing this out. r is Pearson correlation coefficient. We have added this information to the Methods:

"We also calculate the Pearson's correlation coefficient (r) values between the decadal likelihoods and several drought indices for the area. Specifically, we correlate the decadal likelihoods with 10-year running averages of the Palmer Drought Severity Index (PDSI), the Palmer Hydrological Drought Index (PDHI), and the 6-month Standardized Precipitation Index (SP06). Data were downloaded for Oklahoma Climate Division 8 from http://www7.ncdc.noaa.gov/CDO/CDODivisionalSelect.jsp."

RC8) L. 288ff? If gw use is low, what is then used - e.g. for drinking water? Wasn't the whole point to analyse the water source that 'is used'? Very confusing. Citations from 2011 and 2006 should not be cited as 'recent' in this context, as a lot may have changed in 10 years.
AC8. Our wording here was a bit confusing, and we note that we have removed the word "recently" and have tightened it up by saying:

"This is the case because water extraction in the area is relatively low (Christenson et al., 2011) and the groundwater levels are very closely related to rainfall recharge." Please also see response AC3, which relates to this comment.

RC9) L. 292ff The classification into wet and dry decades are nice, but what is the relation to the severity and occurrence of the actual drought events?
AC9. We show in Table 3 the Pearson's correlation coefficient between the groundwater drought probability (i.e., the values in Figure 3) and the 10-year running average of PDSI, PDHI, and SP06. For the benefit of the reviewer, we copy Table 3 below, which shows that the relationship to the agricultural, hydrological, and meteorological indices here is quite strong.

Drought Index		Correlation	
Type	Name	P(Z<Mod)	P(Z<Sev)
Agricultural	Palmer Drought Severity Index (PDSI)	-0.92	-0.83
Hydrological	Palmer Hydrological Drought Index (PDHI)	-0.95	-0.84
Meteorological	Standardized Precipitation Index - 6 month (SP06)	-0.94	-0.82

RC10) How is the link to history made? How can the stakeholder remember what they found important 4 decades ago - this may have been very different from today as life was very different. The constraints on the temporal aspects are not well introduced and not sufficiently discussed.
AC10. We agree that we did not adequately discuss the constraints on the temporal aspects in the previous version. As mentioned in the General response, this is now addressed in point 2 of the conclusions (see General Response and new Figure 9):

"2. Our interviews were conducted following a drought event, and we recognize that the timing of the interviews will likely affect the responses. For instance, interviews conducted during wet or average conditions might elicit less polarized responses, since drought impacts haven't been recently experienced. We note that our approach of applying the interview responses across different climate conditions (i.e., wettest to driest) makes the assumption that the importance of water uses and management preferences are stationary. We acknowledge that different climate conditions, as well as cultural change, technological innovation, climate adaptation, and other processes are likely to influence the cultural factors we investigated here and may mediate how people interact with their environments. Future work could investigate how responses change with different climate conditions over time, and the subsequent implications for drought action. However, hazards and disasters research is almost always conducted immediately after an event, so this is a wise-spread epistemological issue with both pros and cons in terms of what we learn from post-disaster research."

RC11) Section 4. If stakeholders worldview so clearly has opposite rankings in importance, I do not understand why the analysis was carried out on the full sample. Much more logic would be to investigate these two groups separately to obtain more useful results on PDAI or better, incorporate this somehow quantitatively

AC11. Although the stakeholders have different woldviews, and that maps onto their importance rankings, they are still part of the same community and need to make water management decisions together. Part of the purpose of examining the PDAI with all views is to show the spectrum of potential drought action appetite for the community, and how that varies with different water uses.