Reasonably Polarized (Technical Appendix)

Kevin Dorst
kevindorst@pitt.edu
Comments welcome!

Version: 12th September 2020

This is the technical appendix to my blog series *Reasonably Polarized: Why politics is more rational than you think*. It develops the technical details underlying the various arguments, and also addresses further questions and concerns that might come up.

It's a work in progress, and is constantly being revised. Feedback is most appreciated! You can find the most up-to-date version of the appendix here.

Contents

1 Introduction

Here's a link to the full post; it was published on September 5, 2020.

> **Post Synopsis:** I introduced the example of me and Becca, and how we predictably polarized in our political leanings. I argued, using this case, that the standard irrationalist explanation of polarization doesn't work, and sketched how a *rational* explanation would go, previewing the argument to come.
>
> **Appendix Summary:** I don't want to bog down this high-level post with too many details, as those are on their way in the coming weeks. Instead, I'll just provide some relevant links and citations.

Here are some **links to other versions of the core argument this post sets out**:

- 'Why Rational People Polarize' in *Phenomenal World*. This was an early version of the idea that ambiguous evidence can lead to rational polarization.
- 'A Plea for Political Empathy'. This was the opening piece to my *Stranger Apologies* blog (of which this RP-series is a part), on why irrationalist explanations of polarization lead to demonization.
- 'Why the the other side is more rational than you think'. This piece should come out any day now in *Arc Digital*; it develops in more detail the argument that (1) we can't sensibly blame polarization on irrationality, and (2) that we *can* understand it as rationally-caused.

Here are some citations to **articles that endorse the "standard (irrationalist) story" I criticized in the post:**

- Klein (2014), 'How politics makes us stupid'.
- Klein (2020), *Why We're Polarized.*

  Two comments. (1) This is a great book, and I highly recommend it. As we'll see, I buy a good amount of what Klein says when it comes to the empirical story—where I'm skeptical is the way he thinks irrational 'identity-protective cognition' explains polarization.

  (2) Klein *says* he's giving us a rational story: "The American political system... is full of rational actors making rational decisions given the incentives they face. We are a collection of functional parts whose efforts combine into a dysfunctional whole" (xvii). Though I like the sound of that, it turns out that what he means is that individuals are ***pragmatically* rational**—for example, if you really care about your Republican identity, it makes sense to ignore facts that are inconvenient for that identity, because maintaining your identity is more important than getting to the truth. As will become

clearer in Post 4, what I'm interested—and what matters for how we think about the other side—in is whether polarization can be ***epistemically* rational**, i.e. whether people *who care about the truth* can nevertheless wind up predictably polarized.

- Achen and Bartels (2017), *Democracy for Realists*.
- Taber and Lodge (2006), 'Motivated Skepticism in the Evaluation of Political Beliefs'.
- Kahan et al. (2017), 'Motivated Numeracy and Enlightened Self-Government'.
- Nguyen (2018), 'Escape the echo chamber'.

  This is a great piece: it makes some good distinctions between echo chambers and filter bubbles, and it gives a nuanced picture of the (ir)rationality of falling into an echo chamber (or cult). Still, I think it's fair to consider it an irrationalist narrative.
- Lazer et al. (2018), 'The Science of Fake News'.
- Pennycook and Rand (2019), 'Why Do People Fall for Fake News?'
- Van Heuvelen (2007), 'The Internet is making us stupid'.
- Robson (2018), 'The myth of the online echo chamber'.
- Koerth (2019), 'Why Partisans Look At The Same Evidence On Ukraine And See Wildly Different Things'.
- Carmichael (2017), 'Political Polarization Is A Psychology Problem'.

Here are citations to **some other work that's critical of irrationalist explanations of political disagreement:**

- Jern et al. (2014), 'Belief polarization is not always irrational'.
- Benoît and Dubra (2019), 'Apparent Bias: What does attitude polarization show?'
- Singer et al. (2019), 'Rational social and political polarization'.
- Whittlestone (2017), *The importance of making assumptions: why confirmation is not necessarily a bias.*

  See this blog post for an accessible summary of her thesis.
- O'Connor and Weatherall (2018), 'Scientific Polarization'.
- Engber (2018), 'LOL Something Matters'.
- Lepoutre (2020), 'Democratic Group Cognition'.
- Landemore (2017), *Democratic reason: Politics, collective intelligence, and the rule of the many.*

2 How to Polarize Rational People

Here's a link to the full post, published on September 12, 2020.

> **Post Synopsis:** I described an experiment designed to show how it's possible to polarize people using ambiguous evidence. To do this, I introduced *word-completion tasks* in which you're asked to determine whether a given letter-string is completable by an English word. The key point about this task is that it provides *asymmetrically ambiguous evidence*: it's easier to know what to think if there *is* a completion than if there's *not* a completion. As a result, we can split people into groups—the Headsers and the Tailsers—such that Headsers are better at recognizing cases where a coin lands heads, Tailsers are better at recognizing cases where it lands tails, and as a result they predictably polarize.
>
> **Appendix Summary:** Here I'll report in more detail the results of the study I ran confirming this prediction. A pre-registration form for the study is available here.

250 participants were recruited through Prolific (107 F/139 M/4 Other; mean age = 27.06). Subjects were randomly divided into an Ambiguous (A) and Unambiguous (U) condition. Within each condition, they were further (randomly) divided into "Headsers" and "Tailsers". I will abbreviate the groups "**A-Hsers**"; "**A-Tsers**"; "**U-Hsers**", and "**U-Tsers**". Each group was told they'd be given evidence about a series of independent, fair coin tosses.

The A group was informed about how word-completion tasks work, and given three examples ('P_A_ET' [planet], 'CO_R_D', [uncompletable] and '_E_RT' [heart]). The A-Hsers were instructed that they'd see a completable string if the coin landed heads, and an uncompletable if it landed tails. The A-Tsers were instructed vice versa. Each participant was presented with four independent word-completion tasks. In each, they were first told that a coin was flipped to determine (as per the rule above) whether the letter-string they next saw would be completable, and asked how confident they were that it was completable. They used a $0-100\%$ slider to rate this confidence, which they were given standard instructions about how to use. This first ("prior" confidence) question for each toss was an attention check, and participants were instructed to answer "50%" at this stage, since they had not received any evidence. It was pre-registered that I would exclude data from participants who failed two or more of these attention checks. (All in all, data from 25 subjects out of 250 were excluded for these reasons.)

After each check, the participant was presented with some evidence about the coin toss. Of the four tosses each participant saw, two landed heads and two landed tail, so each saw two completable strings and two uncompletable strings, in random orders. The completable strings were randomly drawn from the list, {FO_E_T, ST_ _N, FR_ _L} (forest/foment; stain/stern; frail/frill); the uncompletable strings were drawn from the list, {TR_P_R, ST_ _RE, P_G_ER}.[1] After seeing their string for 7 seconds, the participants were asked how confident they were that it was completable, and presented with a slider between 0–100%.[2]

The U group, in contrast, was told that each toss of the coin would be used to determine the contents of the urn. For U-Hsers, if the coin landed heads then the urn contained 1 black marble and 1 non-black marble; if it landed tails, it contained two non-black marbles. (For U-Tsers, 'heads' and 'tails' were reversed.) The colors of the non-black marbles changed across trials to make clear they were different urns. Again, each toss started by telling them a new coin had been tossed, and asking how confident they were that it landed heads (U-Tsers: how confident they were that it landed tails). This was an attention check as those above; they were instructed to answer "50%", and data from subjects who failed two or more was omitted. Subjects were then told what color marble came from a single random draw of the urn, and asked how confident they were that the coin landed heads (U-Tsers: tails). Each subject saw four separate coin-toss/urn pairs; three of the four revealed a non-black marble, while the fourth revealed a black one—simulating the expected rate of black marbles if the coin landed heads/tails 50% of the time, and the marbles were drawn at their expected rate.

The U group was so designed in order to test the hypothesis that it is *ambiguity* of evidence that drives the polarization effect. As can be seen, there is a structural similarity but also a structural dissimilarity between the A-group and the U-group. The similarity is that both groups have some chance of getting decisive evidence in favor of a hypothesis (finding a completion for the A group; seeing a black marble for the U-group), and some chance of getting weak evidence against that hypothesis (failing to find a completion for the A group; seeing a non-black marble for the U-group).

The dissimilarity is that subjects in the U-group are, in principle, able to know what they should do with this evidence—a straightforward Bayesian calculation says that if you're a U-Hser and see a black marble, you should assign probability 1 to the coin landing heads, and if you see a non-black marble, you should assign probability $\frac{1}{3}$ to it having landed heads.[3] In contrast, with the word completion task when you

[1] No doubt it would be useful to run a study using a bigger sample of letter-strings; obviously, they must be chosen with some care, as completely random strings like 'X_TNO_' will standardly be too obvious.

[2] Pilot studies indicated that when people were asked how confident they were that the coin landed *heads* or *tails*, subjects were substantially confused, often—it seems, reversing the scale or not moving it from 50%. This makes sense, as there is a pragmatic oddness and an extra cognitive load in asking about whether the coin landed heads (tails), when that is known to be equivalent to whether the string was completable. For this reason, I elicited their opinion about whether the string was completable, and used that reported number to infer their confidence in heads/tails on the given flip, based on which group they were in.

[3] $P(H|\textit{non-black}) = \frac{P(\textit{non-black}|H)\cdot P(H)}{P(\textit{non-black}|H)\cdot P(H)+P(\textit{non-black}|\neg H)\cdot P(\neg H)} = \frac{0.5\cdot 0.5}{0.5\cdot 0.5+0.5\cdot 1} = \frac{1}{3}$.

do find a word, you should of course be certain that it's completable; but when you *don't* find a word, it is much harder to know what opinion you should have. This is a theoretical prediction we'll go more into after we get a theory of ambiguous evidence on the table—but the basic idea is that you should be unsure whether you *should* think of a completion, and therefore unsure what evidence you actually received—unsure whether you should be sure the string is completable. (Staring at '_EAR_T', you think, "I don't see one; but should I?" When 'learnt' pops into you're head, you may well think, "Ah, I should've seen that!") Because of this uncertainty, subjects will (rationally) be unsure how much to lower their confidence that the string is completable.

From the responses of each group to each question, I calculated their prior and posterior confidence that the coin landed heads in each toss (for Hsers, this was the number they reported as their confidence; for Tsers, it was obtained by subtracting this number from 100). It was predicted (predictions 1–3) that the ambiguous evidence would lead to polarization, and (predictions 4–6) that it would lead to *more* polarization than the unambiguous evidence:

1. The mean A-Hser posterior in heads would be higher than the prior (of 50%).
2. The mean A-Tser posterior in heads would be lower than the prior (of 50%).
3. The mean A-Hser posterior would be higher than the mean A-Tser posterior in heads.
4. The mean A-Hser posterior would be higher than the mean U-Hser posterior.
5. The mean A-Tser posterior would be lower than the mean U-Tser posterior.
6. The mean difference between A-Hser posteriors and A-Tser posteriors would be larger than that between the U-Hser posteriors and U-Tser posteriors.

Predictions 1, 2, 3, 5, and 6 were confirmed with statistically significant results; Prediction 4 had the divergence in the correct direction but it was not statistically significant. Plots of prior and posterior mean confidences in each group (by-item), along with 95% confidence intervals, displayed in Figure 1:

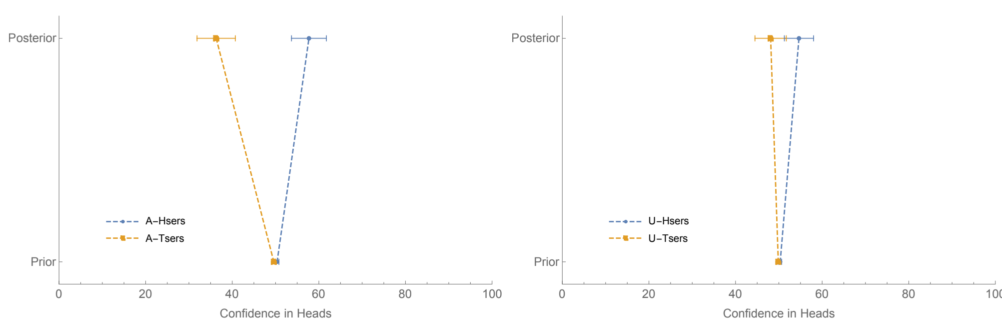

Figure 1: Mean prior and posterior confidence in heads in ambiguous- (left) and unambiguous-evidence (right) conditions. Bars represent 95% confidence intervals.

In more detail: one-sided paired t-test for Prediction 1 indicated that A-Hser priors (M = 50.35, SD = 3.26) were lower than A-Hser posteriors (M = 57.71, SD = 30.33) with $t(219) = 3.58$, $p < 0.001$, $d = 0.341$. One-sided paired t-test for Prediction 2 indicated that A-Tser posteriors (M = 36.29, SD = 31.04) were lower than A-Tser priors (M = 49.60, SD = 2.90), with $t(191) = 5.90$, $p < 0.001$, $d = 0.604$. And one-sided independent samples t-test for Prediction 3 indicated that A-Hser posteriors (M = 57.71, SD = 30.33) were higher than A-Tser posteriors (M = 36.29, SD = 31.04), with $t(410) = 7.07$, $p < 0.001$, $d = 0.699$. Meanwhile, one-sided independent samples t-test for Prediction 4 failed to indicate that A-Hser posteriors (M = 57.71, SD = 30.33) were higher than U-Hser posteriors (M = 54.64, SD = 26.93), with $t(441) = 1.15$, $p = 0.125$, $d = 0.107$. But one-sided independent samples t-test for Prediction 5 indicated that U-Tser posteriors (M =48.10, $SD = 28.47$) were above A-Tser posteriors (M = 36.29, SD = 31.04), with $t(393) = 4.07$, $p < 0.001$, $d = 0.398$.

Prediction 6 was (due to my oversight) handled poorly at the pre-registration stage—I only planned to calculate 95% confidence intervals for the differences between A-Hser and A-Tser posteriors as well as U-Hser and U-Tser posteriors, and compare them. This comparison went as expected: the 95% confidence interval for the difference between A-Hsers and A-Tsers was [15.2, 27.2], while that for the difference between U-Hsers and U-Tsers was [1.8, 11.8]. The former dominates the latter, indicating a larger difference.

What *should've* been planned, I later realized, was to do (a) a 2 × 2 ANOVA, and (b) an empirically bootstrapped 95% confidence interval for the *difference* between the differences between A-Hsers/A-Tsers and U-Hsers/U-Tsers.

(a) Let **valence** be the variable for whether the subject was a Headser (= 1) or Tailser (= 0), and **ambiguity** be the variable for whether the subject was in the ambiguous (= 1) or unambiguous (= 0) group. Analyzing the results using a 2 (valence: Headser vs. Tailser) by 2 (ambiguity: ambiguous vs. unambiguous) ANOVA indicated that there was a main effect of valence ($F(1, 899) = 46.47$, $p < 0.001$), a main effect of ambiguity ($F(1, 899) = 4.31$, $p = 0.038$), and (as should've been predicted) an interaction effect between valence and ambiguity ($F(1, 899) = 14.57$, $p < 0.001$), indicating that the divergence between Headsers and Tailsers was exacerbated by having ambiguous evidence.

(b) Meanwhile, the empirically bootstrapped 95% confidence interval for the difference between differences between A-Hsers/A-Tsers and U-Hsers/U-Tsers was [7.2, 22.6], indicating that the Hsers and Tsers in the ambiguous condition diverged in opinion more than in the unambiguous condition. As mentioned, this condition had a Cohen's d effect size of 0.699. Notably, there *was* a significant difference between U-Hser posteriors (M = 54.64, SD = 26.93) and U-Tser posteriors (M = 48.10, SD = 28.47), with $t(486) = 2.61$ and (two-sided) $p = 0.009$, but the effect size was smaller ($d = 0.236$).

A further oversight on my part at the pre-registration phase was that I only realized after the fact that I actually had access to **time-series data** about how the participant's confidence evolves over time. In particular, using their priors and posteriors for each of the four coin tosses, I could calculate their average confidence in heads after seeing n bits of evidence, for n ranging from 0 to 4.[4] If they are Bayesian in their confidence, this average confidence equals their estimate for the proportion of times the coin landed heads.[5] In fact, using these numbers (and assuming they treated each coin flip independently, as instructed) we could calculate how their opinions would've evolved in any proposition about the coin-flips if they were Bayesian.

Thus I was able to track how their estimate of the proportion of heads (as well as other measures of their beliefs about heads) evolved over time. In other words, we can re-run the above data by pooling responses within subjects and calculating them at each stage in their progression through the experiment. All the predicted results above hold true with this way of carving up the data (with universally lower p-values, since the variance of the data has dropped since we've pooled data within subjects; Prediction 5 is still the only non-significant effect).

Using this, we can calculate the trajectories of their the mean estimate of the proportion of heads (i.e. calculate the mean of the subjects' average confidence in heads at each stage in the experiment), as reported in the blog post; see Figure 2 for this evolution in both the A and U groups.

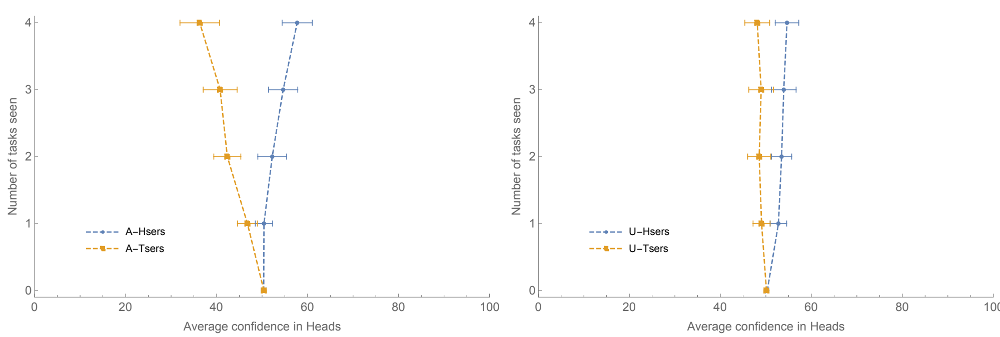

Figure 2: Mean confidence in heads trajectories as subjects view more tasks. Left is ambiguous condition, right is unambiguous condition. Bars represent 95% confidence intervals.

As can be seen, the ambiguous group continues to diverge, while the unambiguous group does not. Using this pooled-within-subject data, a one-sided independent samples t-test indicated that there is an even larger difference between A-Hser posteriors ($M = 57.71, SD = 12.26$) and A-Tser posteriors ($M = 36.29, SD = 14.95$), with $t(101) = 7.98$,

[4] I.e. at stage 0 average their priors for all tosses; at stage 1, average their posterior for the first toss they saw with their priors from the 3 remaining; at stage 2, average their posteriors for the first two tosses they saw with their priors from the remaining 2, etc.

[5] Where P if their probabilistic credence function and I_{H_i} is the indicator variable for H_i (1 if heads, 0 if tails), $\sum_{i=0}^{4} \frac{P(H_i)}{4} = \sum_{i=0}^{4} \frac{\mathbb{E}(I_{H_i})}{4} = \mathbb{E}[\sum_{i=1}^{4} \frac{I_{H_i}}{4}] = \mathbb{E}[\textit{proportion of heads}]$.

$p < 0.001$, and $d = 1.577$. The 95% confidence interval for the difference is now [16.02, 26.82]. (Meanwhile, a two-sided t-test for the difference between U-Hser posteriors ($M = 54.64, SD = 10.19$) and U-Tser posteriors ($M = 48.10, SD = 10.53$) was again significant, but again with a smaller effect size: $t(120) = 3.49$, $p = 0.0034$, $d = 0.631$. The 95% confidence interval for the difference between U-Hsers and U-Tsers was [2.82, 10.26]—again dominated by the A-group's difference confidence interval; likewise, a 2x2 ANOVA again indicated significant main and interaction effects; etc.)

This time-series data allows us to see the divergence in other ways. For instance, consider claims of the form,"there were at least $\boldsymbol{x}$ heads", for x ranging from 1 to 4, and we see the diverging trajectories in the ambiguous condition on the left of Figure 3 (page 10), along with the smaller divergences of the unambiguous condition on the right.

The crucial question: what drives the polarization? The full theory of this will have to wait till we get the theory of ambiguous evidence on the table. But we can start by getting a few things on the table.

First, **the effect is *not* being driven by an asymmetry between strong and weak evidence**. The U group was set up to mirror this asymmetry, but they did not display nearly as strong a polarization effect. Moreover, there is reason to expect that the effect they *did* experience was a "response bias" effect.[6]

Rather, what drives the effect has to do with the *ambiguity* of the evidence—the fact that it's hard to know what to think when you don't find a completion than when you do find one. There is evidence for this in the data. For example, we can divide the A-group cases in which they found a word (operationalized as: they had credence 100 that there was a completion; more on this in a few weeks) and didn't, and then calculate the expected variance in opinion (by weighting the two by what proportion of trials found word) if there *is* a word compared to when there isn't. Likewise, we can divide the U-group into cases in which they saw a black marble vs. didn't, and then calculate the expected variance in opinion if there *is* a black marble vs. if there isn't. The fact that there is an "ambiguity asymmetry" should mean that the expected variance in opinion is asymmetric around the completable/not-completable distinction, but that it is not asymmetric around the black-marble/no-black-marble distinction.

This is what we find. For the A-group: the expected variance in opinion when a word-completion task is completable was 383.9, whereas when it *wasn't* completable it was 746.7. In contrast, for the U-group: the expected variance in a opinion when there was a black marble was 239.0, whereas when there *wasn't* was 266.1.

[6] U-Hsers were asked "How confident are you that the coin landed heads?", whereas U-Tsers were asked "How confident are you that the coin landed tails?". (This mirrors the fact that A-Hsers and A-Tsers were both asked how confident they were that there was a completion—for the former, that question is equivalent to asking for their confidence in heads; for the latter, it's equivalent to asking for their confidence in tails.) It seems probable that this difference in questions drove what effect there was.

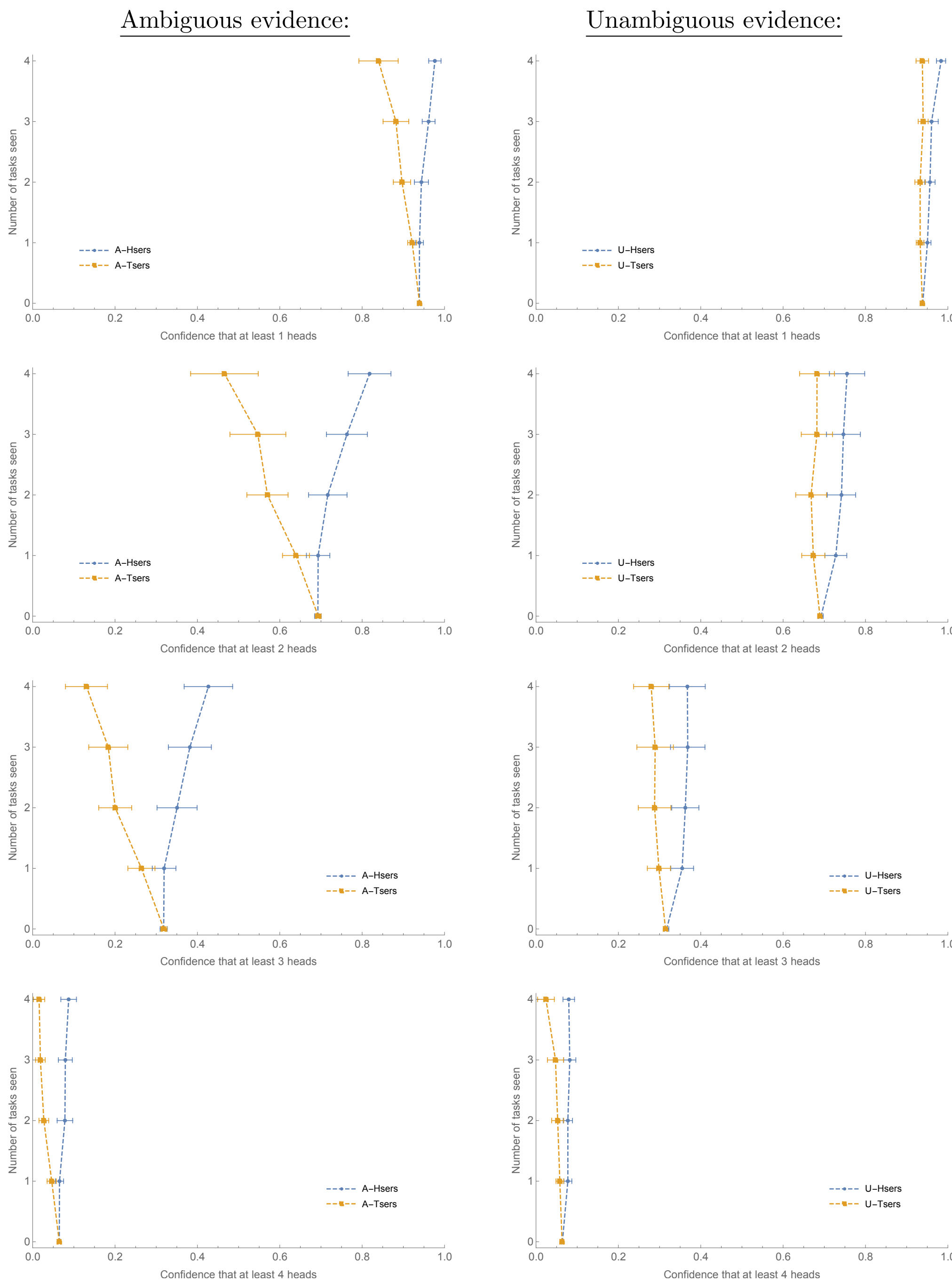

Figure 3: Mean confidence in "at least x heads" as view more tasks, for x ranging from 1–4. Left side is ambiguous condition; right side is unambiguous condition.

I'll argue, in future posts, that this asymmetry in ambiguity in the A-group case is what drives the polarizing effect, for it makes Headsers better at recognizing heads-cases, and Tailsers better at recognizing tails-cases. For now, we can simply report that the data fits with this explanation. When we pool across subjects and divide trials into heads-cases and tails-cases, here are the average posteriors in the groups:

Average confidence it landed heads across cases:

	A-Hsers	A-Tsers	U-Hsers	U-Tsers
Overall:	57.7	36.29	54.64	48.10
Heads cases:	67.42	47.73	66.89	59.95
Tails cases:	48.00	24.84	42.39	36.25

As can be seen, Headsers are better at recognizing heads-cases, Tailsers are better at recognizing tails-cases—and these differences are especially stark amongst the A group. After the fact, I decided to look at these differences statistically (so these tests were not pre-registered). A one-sided t-test found that amongst heads cases, A-Hsers ($M = 67.42, SD = 30.38$) had a mean significantly above A-Tsers ($M = 47.74, SD = 27.54$), with $t(204) = 4.84$, $p < 0.001$, $d = 0.677$. (A-Tsers mean confidence of 47.74 is not significantly different from 50: two-sided t-test yielded $t(95) = -0.80$, $p = 0.423$.) Meanwhile, amongst all tails-cases, A-Hsers ($M = 48.00, SD = 27.10$) were again significantly higher than A-Tsers ($M = 24.84, SD = 30.22$), with $t(204) = 5.80$, $p < 0.001$, $d = 0.810$. (A-Hsers mean confidence of 48.00 is not significantly different from 50: two-sided t-test yielded $t(109) = -0.77$, $p = 0.440$.)

In contrast, these asymmetries were smaller for the unambiguous condition. The difference between U-Hsers ($M = 66.89, SD = 30.27$) and U-Tsers ($M = 59.95, SD = 17.16$) in heads-cases was significant but small: two-sided t-test revealed $t(195) = 2.21$, $p = 0.028$, $d = 0.281$. Meanwhile the difference in tails-cases between U-Hsers ($M = 42.39, SD = 15.43$) and U-Tsers ($M = 36.25, SD = 32.42$) was not significant: $t(169) = 1.88$, $p = 0.062$.

References

Abrams, Samuel J. and Fiorina, Morris P., 2012. 'The big sort that wasn't: A skeptical reexamination'. *PS - Political Science and Politics*, 45(2):203–210.

Achen, Christopher H and Bartels, Larry M, 2017. *Democracy for realists: Why elections do not produce responsive government*, volume 4. Princeton University Press.

Benoît, Jean Pierre and Dubra, Juan, 2019. 'Apparent Bias: What Does Attitude Polarization Show?' *International Economic Review*, 60(4):1675–1703.

Bishop, Bill, 2009. *The big sort: Why the clustering of like-minded America is tearing us apart*. Houghton Mifflin Harcourt.

Brownstein, Ronald, 2016. 'How the Election Revealed the Divide Between City and Country'. *The Atlantic*.

Carmichael, Chloe, 2017. 'Political Polarization Is A Psychology Problem'.

Engber, Daniel, 2018. 'LOL something matters'. *Slate*, 8:1–18.

Fiorina, Morris P, 2016. 'The Political Parties Have Sorted'. *Hoover Institute*, 2(2):1–20.

Gao, George and Smith, Samantha, 2016. 'Presidential job approval ratings from Ike to Obama'. Technical report.

Iyengar, Shanto, Lelkes, Yphtach, Levendusky, Matthew, Malhotra, Neil, and Westwood, Sean J., 2019. 'The origins and consequences of affective polarization in the United States'. *Annual Review of Political Science*, 22:129–146.

Jern, Alan, Chang, Kai Min K., and Kemp, Charles, 2014. 'Belief polarization is not always irrational'. *Psychological Review*, 121(2):206–224.

Kahan, Dan M., Peters, Ellen, Dawson, Erica Cantrell, and Slovic, Paul, 2017. 'Motivated numeracy and enlightened self-government'. *Behavioural Public Policy*, 1:54–86.

Klein, Ezra, 2014. 'How politics makes us stupid'. *Vox*, 1–14.

———, 2020. *Why We're Polarized*. Profile Books.

Koerth, Maggie, 2019. 'Why Partisans Look At The Same Evidence On Ukraine And See Wildly Different Things'. *FiveThirtyEight*.

Landemore, Hélène, 2017. *Democratic reason: Politics, collective intelligence, and the rule of the many*. Princeton University Press.

Lazer, David, Baum, Matthew, Benkler, Jochai, Berinsky, Adam, Greenhill, Kelly, Metzger, Miriam, Nyhan, Brendan, Pennycook, G., Rothschild, David, Sunstein, Cass, Thorson, Emily, Watts, Duncan, and Zittrain, Jonathan, 2018. 'The science of fake news'. *Science*, 359(6380):1094–1096.

Lepoutre, Maxime, 2020. 'Democratic Group Cognition'. *Philosophy & Public Affairs*, 48(1):40–78.

Nguyen, C. Thi, 2018. 'Escape the echo chamber'. *Aeon*.

O'Connor, Cailin and Weatherall, James Owen, 2018. 'Scientific Polarization'. *European Journal for Philosophy of Science*, 8(3):855–875.

Pennycook, By Gordon and Rand, David, 2019. 'Why Do People Fall for Fake News?

Are they blinded by their political passions? Or are they just intellectually lazy?'

Robson, David, 2018. 'The myth of the online echo chamber'.

Singer, Daniel J, Bramson, Aaron, Grim, Patrick, Holman, Bennett, Jung, Jiin, Kovaka, Karen, Ranginani, Anika, and Berger, William J, 2019. 'Rational social and political polarization'. *Philosophical Studies*, 176(9):2243–2267.

Taber, Charles S and Lodge, Milton, 2006. 'Motivated Skepticism in the Evaluation of Political Beliefs'. *American Journal of Political Science*, 50(3):755–769.

Van Heuvelen, Ben, 2007. 'The Internet is Making us Stupid'. *Salon*.

Whittlestone, Jess, 2017. 'The importance of making assumptions : why confirmation is not necessarily a bias'. (July).