Parameters

* timestep = 10 ms
* stateStep = .01
* simCutoff = 20000
* simCount = 8
* seed = 4

Progress: When did the best fitting estimates match the true parameter?

What index was the best fitting estimates? How many at that index?

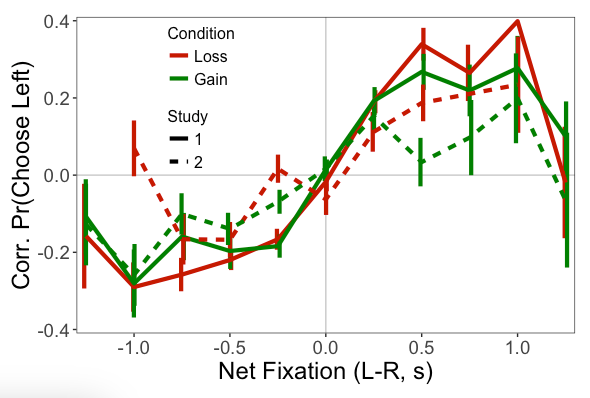
* + 36 mins
  + Gain
    - 1st: 5
    - 2nd: 3
    - Other: 0
  + Loss
    - 1st: 5
    - 2nd: 1
    - Other: 10 (d theta off by 1 unit), 4 (theta off by 2 units)
  + 52 mins
  + Gain
    - 1st: 4
    - 2nd: 2
    - Other: 3 (bias off by 1 unit), 13 (theta off by 3 units)
  + Loss
    - 1st: 6
    - 2nd: 1
    - Other: 44 (sigma theta off by 1 unit)
  + 45 mins
  + Gain
    - 1st: 3
    - 2nd: 2
    - Other: 19 (eta sigma off 1 unit), 5 (eta sigma off 1 unit), 5 (d eta off 1 unit)
  + Loss
    - 1st: 4
    - 2nd: 1
    - Other: 7 (eta sigma off 1 unit), 12 (eta sigma off 1 unit), 3 (eta off 1 unit)
  + 198 mins
  + Gain
    - 1st: 1
    - 2nd: 3
    - Other: 10 (bias off 1 unit), 4 (eta off 1 unit), 12 (bias eta off 1 unit), 8 (bias off 1 unit; eta off 2 units)
  + Loss
    - 1st: 5
    - 2nd: 0
    - Other: 9 (sigma off 1 unit), 5 (eta sigma off 1 unit), 9 (eta sigma off 1 unit)
  + 401 mins
  + Gain
    - 1st: 0
    - 2nd: 2
    - Other: 32, 447, 3, 58, 42, 3
  + Loss
    - 1st: 0
    - 2nd: 0
    - Other: 604, 237, 92, 31, 59, 26, 97, 15
  + We should not be combining the additive and multiplicative model.
* m=1,-6
  + 38 mins
  + Gain
    - 1st: 5
    - 2nd: 0
    - Other: 3 (theta off 1 unit), 3 (minValue wrong; theta off 1 unit), 3 (theta off 1 unit)
  + Loss
    - 1st: 3
    - 2nd: 0
    - Other: 3 (minValue wrong; theta off 1 unit), 4 (minValue wrong; theta off 1 unit), 3 (theta off 1 unit), 3 (theta off 1 unit), 3 (theta off 1 unit)
  + In losses, minValue is typically wrong when there is close to no attentional bias and drift rate is small.
* m=0,-7
  + 38 mins
  + Gain
    - 1st: 4
    - 2nd: 1
    - Other: 4 (t off 1 unit), 4 (t off 1 unit), 4 (s off 1 unit)
  + Loss
    - 1st: 4
    - 2nd: 4
    - Other:
  + In losses, minValue is typically wrong when there is close to no attentional bias. This makes sense since the impact of minValue goes to 0 as theta goes to 1.
  + 78 mins
  + Gain
    - 1st: 4
    - 2nd: 1
    - Other: 3 (minValue wrong), 6 (minValue wrong; theta off 1 unit), 6 (minValue wrong; theta off 1 unit)
  + Loss
    - 1st: 3
    - 2nd: 0
    - Other: 3 (minValue wrong; theta off 1 unit), 4 (theta off 1 unit), 3 (theta off 1 unit), 3 (theta off 1 unit), 3 (theta off 1 unit)
* GEN: ; FIT:
  + 52 mins
  + Gain
    - 8/8 Best fitting model was always dstm
    - 1st: 5
    - 2nd: 0
    - Other: 3 (theta off 1 unit), 3 (minValue wrong; theta off 1 unit), 4 (theta off 1 unit)
  + Loss
    - 8/8 Best fitting model was always dstm
    - 1st: 3
    - 2nd: 0
    - Other: 4 (minValue wrong; theta off 1 unit), 5 (minValue; theta off 1 unit), 3 (theta off 1 unit), 5 (theta off 1 unit), 3 (theta off 1 unit)
* GEN: ; FIT:
  + 52 mins
  + Gain
    - 7/8 Best fitting model was dse, 1/8 dse was second best
    - 1st: 2
    - 2nd: 3
    - Other: 19 (eta sigma off 1 unit), 5 (eta sigma off 1 unit), 5 (d eta off 1 unit)
  + Loss
    - 7/8 Best fitting model was dse, 1/8 dse was third best
    - 1st: 3
    - 2nd: 1
    - Other: 7 (eta sigma off 1 unit), 12 (eta sigma off 1 unit), 3 (wrong model; d off by 1 unit, eta and theta wrong… though true eta=.004 theta=1, est eta=0 theta=.5), 3 (eta off 1 unit)
* GEN: ; FIT:
  + 52 mins
  + Gain
    - 8/8 Best fitting model was dstmr
    - 1st: 3
    - 2nd: 1
    - Other: 3 (minValue wrong), 6 (minValue wrong; theta off 1 unit), 3 (range wrong; drift off 2 units; data generating model was 3rd best), 7 (minValue wrong; theta off by 1)
  + Loss
    - 8/8 Best fitting model was dstmr
    - 1st: 1
    - 2nd: 2
    - Other: 6 (minValue range wrong; t off 1 unit), 8 (t off 1 unit), 6 (t off 1 unit), 8 (t off 1 unit), 6(t off 1 unit)
  + dse was never the best fitting model, which is a good sign. But dstmr has problems when there is little or a lot of attentional bias.
  + When theta is close to 1, the model has a hard time telling the minimum value since it matters less (i.e. the value difference ends up being the same regardless of how we transform values). This is true for dstm too.
  + When theta is close to 1 or 0, range and drift can trade off. Writing the model with d=d and r=5r is the same as writing it with d=1/5d and r=r.
* GEN: ; FIT:
  + 52 mins
  + Gain
    - 7/8 Best fitting model was dse
    - 1st: 2
    - 2nd: 2
    - Other: 19 (eta sig off 1 unit), 3 (dst with slight attn bias was best fit), 5 (eta sig off 1 unit), 5 (d eta off 1 unit)
  + Loss
    - 7/8 Best fitting model was dste
    - 1st: 3
    - 2nd: 1
    - Other: 7 (d off 1 unit; eta sig off 2 units), 12 (eta sig off 1 unit), 5 (dst best fitting), 3 (eta off 1 unit)
* GEN: dst; FIT: dstm + dst
  + Gain
    - 7/8 Best fitting model was dst
    - 1st: 4
    - 2nd: 1
    - Other: 5 (t off 1 unit), 4 (t off 1 unit), 5 (s off 1 unit)
  + Loss
    - 8/8 Best fitting model was dst
    - 1st: 5
    - 2nd: 1
    - Other: 10 (d t off 1 unit), 4 (t off 1 unit)
* GEN: dstm; FIT: dstm + dst
  + Gain
    - 6/8 Best fitting model was dstm
    - 1st: 4
    - 2nd: 1
    - Other: 3 (minValue wrong), 5 (minValue wrong; t off 1 unit), 4 (minValue wrong; t off 1 unit)
  + Loss
    - 7/8 Best fitting model was dstm (multiple best fitting when theta estimated to be 1)
    - 1st: 3
    - 2nd: 0
    - Other: 4 (minValue wrong; t off 1 unit), 5 (minValue wrong; t off 1 unit), 3 (t off 1 unit), 5 (t off 1 unit), 3 (t off 1 unit)
  + When theta estimated to be 1, minValue can take on multiple values at same NLL.
* GEN: dstmr; FIT: dstmr + dst
  + Gain
    - 8/8 Best fitting model was dstmr (multiple best fitting)
    - 1st: 3
    - 2nd: 1
    - Other: 4, 8, 3, 7
  + Loss
    - 8/8 Best fitting model was dstmr (multiple best fitting)
    - 1st: 1
    - 2nd: 2
    - Other: 7, 9, 6, 7, 6
  + dstmr struggles to differentiate from dst when theta close to 1. Drift is way off in these cases.
  + Same NLL for different ranges if we adjust drift.
* GEN: dst; FIT: dstmr + dst
  + Gain
    - 7/8 Best fitting model was dst (multiple best fitting)
    - 1st: 2
    - 2nd: 2
    - Other: 8, 3, 7, 8
  + Loss
    - 8/8 Best fitting model was dst
    - 1st: 5
    - 2nd: 1
    - Other: 10, 4
* GEN: dstmr; FIT: dstmr + dstm
  + Gain
    - 8/8 Best fitting model was dstmr (multiple best fitting)
    - 1st: 5
    - 2nd: 0
    - Other: 4 (minValue wrong), 9 (minValue wrong; t off 1 unit), 9 (minValue wrong; t off 1 unit)
  + Loss
    - 8/8 Best fitting model was dstmr (multiple best fitting)
    - 1st: 2
    - 2nd: 1
    - Other: 8 (minValue wrong; t off 1 unit), 10 (t off 1 unit), 9 (t off 1 unit), 9 (t off 1 unit), 9 (t off 1 unit)
* GEN: dstm; FIT: dstmr + dstm
  + Gain
    - 8/8 Best fitting model was dstm (multiple best fitting)
    - 1st: 3
    - 2nd: 2
    - Other: 7 (t off 1 unit), 7 (t off 1 unit), 7 (t off 1 unit)
  + Loss
    - 8/8 Best fitting model was dstm (multiple best fitting)
    - 1st: 2
    - 2nd: 1
    - Other: 4 (minValue wrong; t off 1 unit), 11 (t off 1 unit), 7 (t off 1 unit), 7 (t off 1 unit), 7 (t off 1 unit)

**What did I learn from parameter recovery?**

* Now we know that multiplicative and additive models can be told apart when the data generating process is multiplicative or additive. It’s a little bit trickier differentiating between the different types of multiplicative models; for instance, we might find evidence of range-normalization in gains when there is none. This would be with the same likelihood as a model with no range normalization but a smaller drift rate (see next point).
* Range normalized aDDM runs into an identification problem. The range can be factored out of the drift and multiplied with the drift rate parameter. This means and will give the same NLL as and . This means *we can’t identify whether range normalization or goal relevant is occurring, since the drift rate can change and result in equal likelihoods.* We need some further assumption to rule out the two identical models if an RNaDDM or its equivalent is the best-fitting model.
* Any goal-relevant model that looks at value deviations from some reference point (e.g. the minimum value in a condition) is going to struggle to identify if this is occurring when theta is close to 1. This is because if you rearrange the aDDM to make minValue additively separable from and , it’s effect on the choice process converges to 0 as theta goes to 1.
* The goal-relevant model also struggles to fit to our gain condition in the numeric task since the reference point is 1, which is not far from 0. I have a feeling that when people are choosing in gains, they’re just using 0 as a reference point. In losses, if they’re accumulating positive evidence when looking at an aversive lottery, then this reference point can be the minimum value, or perhaps even something lower than that (e.g. a lower round number, like minValue-1).
* If theta<1, then goal-relevant models counteract attentional biases in gains (i.e. when minValue>0) and amplify attentional biases in losses (i.e. when minValue<0). Again, you can see this when you make minValue additively separable from raw value signals. It’s like making attentional effects weaker in gains than in losses. This is not true in our data.
  + 

**What do I think the next steps should be?**

* Change the narrative from goal-relevant evidence to reference-dependent evidence. Try parameter recovery with 2 more reference-dependent models:
  + A round number below the minimum value in a context. Say 0 in gains and -7 in losses. If (1,-6) is distinguishable from additive model, then I think we’re ok just checking that we can recover the OG data generating process for itself and not do pairwise comparison. Very good recovery (see summary above). I think we should try this in the data too.
  + the average value in a context (if their attentional choice biases reflect this… check that first). THIS CAN’T BE THE CASE BECAUSE THERE’S STILL NET FIXATION BIAS IN TRIALS WHERE BOTH VALUES ARE BELOW AVERAGE IN THAT BLOCK. SEE PLOT BELOW.



* We need to decide if we’re going with range normalized or reference dependent, but I don’t think we can test both at the same time. They’ll pop out equal NLLs with different drift rates. I think we should go with reference dependent, since it builds on Sepulveda et al (2021) and range normalization could conflict in the future with Smith and Krajbich (2019) and Ting and Gluth (2023), though the evidence as of right now is not sufficient to argue this. I think it’s a cleaner story to say that people are seeking out redeeming qualities about things they don’t like, than it is to say that people are range normalizing across negative outcomes. I think it also allows the framework to be more flexible in the future, allowing people to test different reference-points.
* Get posterior model probabilities for a few subjects, say 8. See how this goes. I think we can test reference-dependent (1,-6), reference-dependent (0, -7), standard, and additive aDDM.