**Baseline questions:**

I have a few basic questions regarding the results that I think we probably need to write a paper on what we have so far. I have compiled these from both the draft above and the pointers/tables that Batu sent. Feel free to reply to these via email, or since this was a lot to type out, we could set up another meeting with everyone.

1. **Regarding applications included**: **Dr. Oren**, were these submitted through ERAS? If not exclusively, could we detail how they were submitted for the "Methods" section of the paper?
2. **Regarding general methods**: **Batu**, did we make any exclusions to the letters included in the analysis? **All**, should we?

We didnt exclude

1. **Regarding general methods**: **Batu**, did we exclude any outliers (letter writers who used a crazy amount of certain words)? **All**, should we?

We didnt

1. **Regarding general methods**: **Batu**, what statistics software was used (in this format "Stata/SE by RR (version 15; StataCorp LLC; College Station, Texas")?

We used scipy.stats(a python scientific computation library)’s fisher\_exact module to calculate odds ratio and the corresponding p value

*from scipy.stats import fisher\_exact*

1. **Regarding general data**: **Batu**, what are some of these numbers?
   1. # of words submitted for analysis

# of words submitted for analysis is 2105690

* 1. Any letter writers with more than one letter

Since we dont have writer names in our dataset. We can not answer this question

* 1. Average difference of sentences for female versus male writers

Added to the results table

* 1. Average difference of words for female versus male writers

Added to the results table

* 1. Any differences of letter length based on letter rank
  2. Difference of letters submitted for male versus female applicants (likely no significance based on tables so far though)
  3. Difference of sentence and word count for male versus female applicants (likely no significance based on tables so far though)

1. **Regarding IMG analysis**: **All**, I am thinking of removing this component of the paper. While important for the exploratory phase, I am not sure it adds much in the end and is distracting because we do not draw conclusions from it. However, if we do want to keep it in, then, is there a significant difference for the letter writer gender or letter writer rank for IMGs versus non-IMGs?
2. **Regarding applicant's gender**: **Batu**, in the pointers document, it says "Radiology specific descriptors are used more for female applicant than male applicants (p<0.05)" but on the excel tables, the p-value is 0.07. How should I reconcile this? I assumed the excel tables were more up to date.
3. **Regarding writer's gender**: **Batu**, in the pointers document, it says "Female high-rank writers (Full Professors) use Radiology specific descriptors more than male high-rank writers, while male professors predominantly use “standout” language" but on the excel tables, I do not see p-values to indicate significance. Moreover, even if we go on just the numbers/percentages, the N is very small to make a strong conclusion.
4. **Regarding writer's gender**: **All/Batu**, in the pointers document and excel tables, we compare female applicant x female writer to male applicant x female writer, then female applicant x female writer to female applicant x male writer, female applicant x male writer to male applicant x male writer, and male applicant x female writer to male applicant x male writer. This results in your bullet points "Female writers use significantly more (p<0.01) communal language to male applicants than they use to female applicants (counter intuitive in my opinion)" and "Male writers use significantly more (p<0.05) radiology specific language to female applicants than they do to male applicants." This feels redundant and not necessarily statistically meaningful. Do we need to create a multivariate model instead like Grimm et al. where the outcome is the descriptor controlled for by applicant gender, writer gender, writer rank, and applicant gender x writer gender?
5. **Regarding writer's gender**: **Batu**, you say "Female writers use communal language more than male writers (p<0.01)." Can we see what words this was driven by?
6. **Regarding writer's gender and seniority**: **Batu**, in the pointers document you say "Male writers use “standout” language more than female writers (p<0.01), However - this might be due to seniority of the male writers" where I presume more males of higher seniority and letter writers of higher seniority were more likely to use standout terms. Please let me know if I am assuming something wrong because I did not see p-values for the latter, that letter writers of higher seniority were more likely to use standout terms, and based on pure percentage 25% doesn't seem too different from 22% or 21%. Regardless, can we not perform some analysis controlling for this factor to actually prove our hypotheses?
7. **Regarding writer's seniority**: **Batu**, in the demographic chart, you separate the writer's seniority into full professor, associate professor, assistant professor, leadership (dean, chair), and others. What does "others" entail? I also didn't realize we could go so granular as to know if someone is a dean or chair and would like to include that if possible. However, in the charts on the second sheet for word analysis, we drop this category. Can we keep the categories consistent?
8. **Regarding word analysis**: **All**, I am still a little hesitant about the way that we are doing word analysis. I think initially we said these terms were a starting off point and we would become more systematic (e.g., find a dictionary version) as we did more. But at this point, I feel like we are using these words that I kind of pulled from samples that were provided in research papers as an official word list. Even if we end up using it, there is a concept called MECE (mutually exclusive, collectively exhaustive) that I don't think we meet because we have duplicate words across categories. For example, "hard work" is part of both agency and grindstone, so it is not correct I think to generate a p-value off that. Or, why is "industrious" under agency and not grindstone? I just think we need to be more rigorous in this aspect now that we have done some initial analysis. There are also words that we may need to consider adding: mother, wife, father, immig\*, hobby, hobbies, future leader, accomplishment, ability, instinct, expert\*.