Comments to the author (if any):  
Reviewer #2: Overall Review:  
This paper analyzes the transcribed presidential debates between 1980 to 2020 against Moral dimensions. Authors use different moral loading variations to explain contextual changes in the conversations of Democrats and Republican leaders. The authors' one of the key objectives is to understand whether presidential candidates became more focused on their own issue stances and moral reasoning than on bargaining with each other. Moral divergent stance is used in this case. However, the authors draw this in the lens of "mediation/mediatization" theory, and I found this lens disconnected at some point when describing the divergence in morality.  
However, the paper highlights an important phenomenon that computational social scientists may bring impact to the democracies and presidential candidates formulate a logical discussion upfront in their debates. Having said that, the following quotes and comments indicate specific points that I felt the authors need to address in their revisions.  
  
Abstract:  
"Our results show a significant increase in moral divergence between Democrat and Republican candidates since 1980, which implies that presidential candidates became more focused on their own issue stances and moral reasoning than bargaining with each other. " => This statement does not reflect what the study actually did, since it does not explore the operationalising of  bargaining; in my opinion, the authors can not claim the effect of morality than bargaining.  
  
Introduction:  
"Unfortunately, if we use Moral Foundation Theory as a guide (Graham et al., 2013; Haidt & Graham, 2007; Haidt & Joseph, 2004) in conjunction with mediatization theory (Hjarvard, 2008, 2013), we will note that a lack of discussion regarding the primary issues and a tendency to avoid conflict when responding to the issues at hand are inevitable in televised presidential debates" => This whole sentence making the reader bit confusing for a beginning as to understand why is it "unfortunate". I would suggest elaborating more or since this is about the choice of method, be clear on the background information first. Give some context to what could be unfortunate.  
  
"- from a mediatization perspective, and therefore sheds light on the development of ways to encourage more effective political communication" => the contribution of the paper is clear in the moral ground, but how these findings or connection to mediatization does not quite strongly appear in the background provided in the introduction.  
  
Literature review:  
Although I found the Moral foundations and their relationships to political debate are well explained, the connection to Morality and Mediation is not being elaborated. In other words, the hypothesis constructions to resonate with H3 is very weak - or in this context, I still can't figure out the connection. I would suggest the authors to bring a description centering on how morality and mediation have been playing in political dialogs.  
  
Methods:  
The data and methods are articulated well, and easy to follow the process. The authors have well explained the reasonable challenges of using a Moral Foundations Dictionary and how DDR can overcome such challenges.  
  
  
Results  
"Our results show that substantial variances in moral loading occurred at the second level (ICC = .34) and third level (ICC = .17), which indicates that the moral loadings of a given debate on the 10 moral dimensions (5 pairs) are substantially correlated,....". => please indicate more details why intraclass correlation was measured. Also, please indicate the numbers properly with desired numbers (i.e., the range could be from 0 to 1, which indicates the threshold for good correlation).  
  
P.17: "In addition, we examined the moral loading change in each moral dimension. We used the moral loading of each dimension as a dependent variable and year, party and their interaction as independent variables." =>  How did you obtain the interactions? Since you mentioned the unit of analysis is the debate - and I assumed you filtered the democratic candidate speech and republican candidates speech separately, but none of the details are explained.  
  
Discussion:  
"Moreover, by quantitatively exploring how mediatization could have influenced the political debate from a historical perspective, this study overcomes the three main challenges of mediatization research: historicity" => The authors are not doing a reasonable job in explaining how mediation is influencing the political debate. From H3 indicated in the literature, it is not clear how the authors relate such theory to the results.  
  
"Secondly, Democrat candidates' moral loadings are almost always higher than Republican candidates' moral loadings across all five moral foundations." => Is it in virtue or vice that is higher?  
  
"Presidential debaters may respond to moral concerns by elaborating upon their own moral concerns to build their own image rather than engage in or generate real issue discussions, explore possible solutions, or find a way to collaborate with each other."... => As I raised my concerns earlier, I would like to repeat the same question - the moral loading in a speech does not compare anything other than moral dimensions. If they engage in real issues, or talk through to uplift their image- I am not sure if this study has an interpretation or measurement. In other words, if you are to indicate that presidential debates whether discuss real issues - there should have been another method to measure it, rather than the moral loading. If you use moral divergence as a proxy for this, I would suggest providing more context and relevance in the descriptions.  
  
  
  
Reviewer #3: This paper examines US presidential candidates' use of moral language during every televised presidential debate so far. Specifically, it uses word embeddings to detect each candidate's use of language associated with the five moral foundations from Moral Foundations Theory. I believe the manuscript provides descriptive results that will be of use to communication scholars, psychologists,  political scientists, and other students of campaign rhetoric. In particularly, I think it is useful to know how much attention each of the five moral foundations have received from each party over time.  
  
The use of word embeddings strikes me as a better choice than a simple word count based on the Moral Foundation Dictionary. I suspect the use of language has changed too much over this period for a simple word count to produce meaningful over time comparisons. I think the manuscript explains this point clearly and effectively. But the editors should note that I have never used Distributed Dictionary Representations (the specific algorithm used here) myself and thus am not qualified to judge its application in this manuscript.  
  
I have two major concerns, both of which I believe could be addressed in a revised manuscript.  
  
My first concern is that the manuscript's discussion of mediatization seems unnecessary. Large portions of text are dedicated to explaining this concept. Yet as best I can tell,  mediatization is mostly used to justify Hypothesis 3 ("The moral divergence between Democratic [relatively liberal] and Republican  
[relatively conservative] presidential candidates has been increasingly widening."). Yet I can think of many other reasons that the parties would diverge in their use of moral rhetoric (e.g., ideological polarization of the parties, partisan sorting of supporters of each party, the growth of cable news, social media). The research design cannot distinguish between these and other explanations and thus the attention to mediatization seems distracting rather than clarifying. (one could argue that my explanations are each the result of mediatization too, but if so the term seems to broad to provide much theoretical leverage).  
  
My second concern is that the manuscript provides scant details about the text preprocessing. Before analyzing any text, researchers need to make several important choices about the text to input into their models. These choices include decisions about whether to remove stop words, rare words, punctuation, contractions, numbers, and whether to stem the words. These choices often have large consequences for the results of text analysis (see e.g., Matthew J. Denny and Arthur Spirling's 2018 article in \*Political Analysis\*). In this case, I worry that the main conclusions may be sensitive to these choices as well. For instance, it seems that the main analysis relies on the full debate transcript. As a result, the inclusion of stop words may overstate the similarity between Democrats' and Republicans' use of moral rhetoric. Denny and Spirling argue that researchers should provide analysis of the sensitivity of their results to these various choices. Thus, I think an appendix  
with such sensitivity analysis would give the descriptive results far more value.  
  
In addition to those major concerns, I have a few more minor suggestions:  
  
\* I would prefer a regression table displaying the main regression results (perhaps in an appendix). It would have helped me think through the data structure and model and better understand Figure 1.  
\* The measure of difference between the parties seems somewhat arbitrary to me (p. 17: "We added up the absolute value to form a unidimensional score to reflect the total difference between Democrats and Republicans in each debate."). Why not weight each dimension by how much it was used in the debate? By treating each dimension equally, it might miss important differences between the parties on the dimensions receiving the most attention in the debate (or overstate differences due to less important dimensions).  
\* At several points, the text suggests the results reflect the speaker's true morality, but this claim seems unwarranted. For instance, on page 12, the manuscript states, "The result is that presidential debaters may have been increasingly overly focused on their own moral concerns that derived from their own sensitive moral foundations...". Though this \*may\* be true, it strikes me as an overly strong assumption that these changes reflect an increasing reliance on the candidate's own moral concerns, rather than other changes such as the distribution of moral values between supporters of each political party, shifting policy agendas, or the refinement of campaign strategies.  
\* I found several other unjustified claims distracting as well. For instance, the first paragraph of the manuscript ends, "we will note that a lack of discussion regarding the primary issues and a tendency to avoid conflict when responding to the issues at hand are inevitable in televised presidential debates." This manuscript does not demonstrate that it is "inevitable" that candidates will avoid conflict in debates and I can't imagine a research design that could demonstrate this point. It's a minor detail, but I think it represents a broader issue. I encourage the authors to revise the entire manuscript to avoid over-statements such as this and untested claims such as the point above.  
  
In summary, I think this manuscript provides a useful descriptive analysis. It would be even stronger with some revision.