Editorial comments:

Changes to be made by the Author(s) regarding the manuscript:

1. Please take this opportunity to thoroughly proofread the manuscript to ensure that there are no spelling or grammar issues.

We have now thoroughly proofread the revised manuscript.

2. Please revise lines 399-402 to avoid previously published text.

This section of the Discussion has been modified accordingly.

3. Please upload each Figure individually to your Editorial Manager account as a .png, .tiff, .svg, .eps, .psd, .pdf, or .ai file.

We have uploaded all the figures as individual PDF files.

4. Please provide an email address for each author.

We have provided the email address of the corresponding author in the manuscript (Yaoguang Jiang: jiangyaoguang@gmail.com). It is however unclear where in the manuscript we should put the email address of the non-corresponding author (Michael L. Platt: mplatt@pennmedicine.upenn.edu).

5. Please revise the Short Abstract to clearly describe the protocol and its applications in complete sentences between 10-50 words: “Here, we present a protocol to …”

The Short Abstract has been modified accordingly.

6. Please move the ethics statement before your numbered protocol steps, indicating that the protocol follows the animal care guidelines of your institution.

This has been fixed.

7. Please adjust the numbering of the Protocol to follow the JoVE Instructions for Authors. For example, 1 should be followed by 1.1 and then 1.1.1 and 1.1.2 if necessary. Please refrain from using bullets, dashes, or indentations.

We have made sure the numbering system is correct.

8. Please revise the protocol to contain only action items that direct the reader to do something (e.g., “Do this,” “Ensure that,” etc.). The actions should be described in the imperative tense in complete sentences wherever possible. Avoid usage of phrases such as “could be,” “should be,” and “would be” throughout the Protocol. Any text that cannot be written in the imperative tense may be added as a “Note.” Please include all safety procedures and use of hoods, etc. However, notes should be used sparingly and actions should be described in the imperative tense wherever possible.

We have made sure the protocol section conforms to journal standard.

9. Please add more details to your protocol steps. There should be enough detail in each step to supplement the actions seen in the video so that viewers can easily replicate the protocol. Please ensure you answer the "how" question, i.e., how is the step performed? Alternatively, add references to published material specifying how to perform the protocol action. Please see the following examples:

We have inserted the proper citation (Jiang & Platt 2018) for the first experiment, and added more details for the second experiment. However, we want to emphasize that some of these details are not critical for performing the protocols described. For example, the overall number of monkeys in a colony room is irrelevant as long as the monkeys participating in the experiment live in the same colony room (with or without other monkeys) for the duration of the experiment.

1.2: Please specify how many monkeys are included in this experiment.

7 adult macaque monkeys are included in this experiment.

1.3: How many monkeys are there in a colony room?

For the duration of the experiment these 7 monkeys were the only monkeys in the colony room.

2.3: How many different M2s are there? What is the duration for each M2? For how many days is this step repeated?

All 7 participating monkeys have equal probability of being M2. On each day one monkey acts as M1 and faces all the other 6 monkeys, each for 5 minutes. Each treatment condition is repeated 5 times for each M1.

3.1.2, 3.2.2, etc.: For how many days is this step repeated? Is OT or saline given to every monkey or only a specific group? Please specify throughout the protocol.

Treatment is only delivered to M1s. This has been further specified in the protocol.

3.2.2: Is the monkey restrained or anesthetized before injection?

Monkeys are restrained via a head holder fixed on the primate chair. The are not anesthetized.

Line 143: Please describe how to train the monkeys.

This is explicitly explained in the next section (2.2. initial behavior training).

Line 172: How is the eye tracking system installed?

The eye camera is mounted onto the primate chair. We suggest following the instruction provided by the supplier (Eyelink 1000 Plus, SR Research) for installation of the 'primate mount' tracking mode.

Lines 175, 198-205, 213-215: Please describe how.

175: This has been further clarified. Briefly, initial training is done through gradual behavioral shaping through reinforcement.
198-205: Proper references have been added for this procedure.
213-215: We suggest closely following the user manual provided by the supplier (Omni-Plex, Plexon Inc) for these steps.

Lines 211: Is the monkey restrained or anesthetized before performing the penetration?

Monkeys are restrained via a head holder fixed on the primate chair. The are not anesthetized.

10. Please include single-line spaces between all paragraphs, headings, steps, etc.

This has been addressed.

11. After you have made all the recommended changes to your protocol (listed above), please highlight 2.75 pages or less of the Protocol (including headings and spacing) that identifies the essential steps of the protocol for the video, i.e., the steps that should be visualized to tell the most cohesive story of the Protocol.

The relevant sections have been highlighted.

12. Please highlight complete sentences (not parts of sentences). Please ensure that the highlighted part of the step includes at least one action that is written in imperative tense. Please do not highlight any steps describing anesthetization and euthanasia.

This has been noted.

13. Please include all relevant details that are required to perform the step in the highlighting. For example: If step 2.5 is highlighted for filming and the details of how to perform the step are given in steps 2.5.1 and 2.5.2, then the sub-steps where the details are provided must be highlighted.

This has been noted.

14. JoVE articles are focused on the methods and the protocol, thus the discussion should be similarly focused. Please revise the Discussion to explicitly cover the following in detail in 3-6 paragraphs with citations:

a) Critical steps within the protocol
b) Any modifications and troubleshooting of the technique
c) Any limitations of the technique
d) The significance with respect to existing methods
e) Any future applications of the technique

The Discussion has been significantly modified accordingly. We have also incorporated both reviewers' suggestions in this section.

15. References: Please do not abbreviate journal titles.

This has been fixed.

16. Please follow the book citation example below to reformat book references:
Kioh, L.G. et al. Physical Treatment in Psychiatry. Blackwell Scientific Pubs. Boston (1988).

This has been addressed as well.

Reviewers' comments:

Reviewer #1:

This manuscript contains interesting and valuable data but it needs more work before it is acceptable for publication. The manuscript joins (for no clear reasons) two different experiments that have little connection to each other. I believe both datasets warrant a separate publication. The effect of OT and AVP on social dominance, even if relatively weak, is important to document because there is a large species-dependent discrepancy in the literature. The soccer (penalty kick) task also deserves a framework on its own as it is one of the most interesting and potentially useful social tasks that have been recently developed.

We thank the reviewer for their general enthusiasm regarding our behavioral paradigms. While we agree with the reviewer that the bigger scientific questions driving these two studies are different, we do believe the two behavioral paradigms ('face-off' and 'penalty kick') are conceptually related and thus could benefit from being presented together. Specifically, as explained in the Introduction, many existing experimental paradigms in social neuroscience are too rigid and one dimensional, thus failing to capture the interactive, dynamic, and evolving nature of real life social behavior. Both of our paradigms aim to improve upon the existing designs by featuring pairs of monkeys interacting with each other in relatively unconstrained, multi-dimensional decision spaces. The neurobiological investigations accompanying these studies (i.e. how neuropeptides modulate social dominance, and how neural activity in mSTS is correlated with strategic competition) serve more as examples of the potential *applications of these behavioral paradigms, rather than the technical emphasis of the current manuscript. We have now made this point more explicitly in Discussion.*

Both components have good content that would benefit from major revisions. In the following sections, I will treat them separately.

We will try our best to address all of the reviewer's concerns.

The OT/AVP administration experiment
In the presentation of the OT/AVP data, many details are skipped or are unclear. For example, in the results section (third paragraph) the authors should clarify the following: " ….a new colony of 7 monkeys (M1 = 3 males, M2 = 3 males, 4 females)". Does this mean that the same 3 males served both as M1 and M2 but the females were only M2? If yes, this difference has bearing on the results.

The reviewer is correct that in this dataset (OT and AVP inhalation experiment) the male monkeys acted as both M1 and M2s, whereas the females only acted as M2s. By contrast, in the first dataset (OT inhalation only), all 7 male monkeys acted as M1s as well as M2s. We have further clarified this point in the figure legend of the revised manuscript. In the published paper (Jiang & Platt 2018) we have systematically compared these two dataset and determined that including female M2s in the second dataset did not in any way significantly alter the observed behavioral effects of OT inhalation. We want to emphasize that, as far as we know, the effects of OT and AVP inhalation and injection are quite consistent as long as the monkeys receiving the treatments (M1s) are males. In contrast, in a separate manuscript (submitted) we described that the same neuropeptide treatments have very different effects when female macaque monkeys are the ones that receive these treatments.

At some point the notion of "reduction" is not supported by statistical measures. For example, the title of figure 2A- F reads: "OT and AVP inhalation and injection flattens social hierarchy" yet the difference in staring time is $p = 0.098$ (nearly a 10% chance), which indicates that this small difference could be obtained by chance. It would be helpful to show individual data as some monkeys may show the effect whereas others may not. This is particularly important because the effects of OT across many studies from many laboratories are small and may not manifest in all subject animals.

The reviewer's point is well understood. It is absolutely the case that the behavioral effects of neuropeptides are not consistent across individual monkeys, as reported in other animal and human research (for example see Rebuttal Figures 1-2).

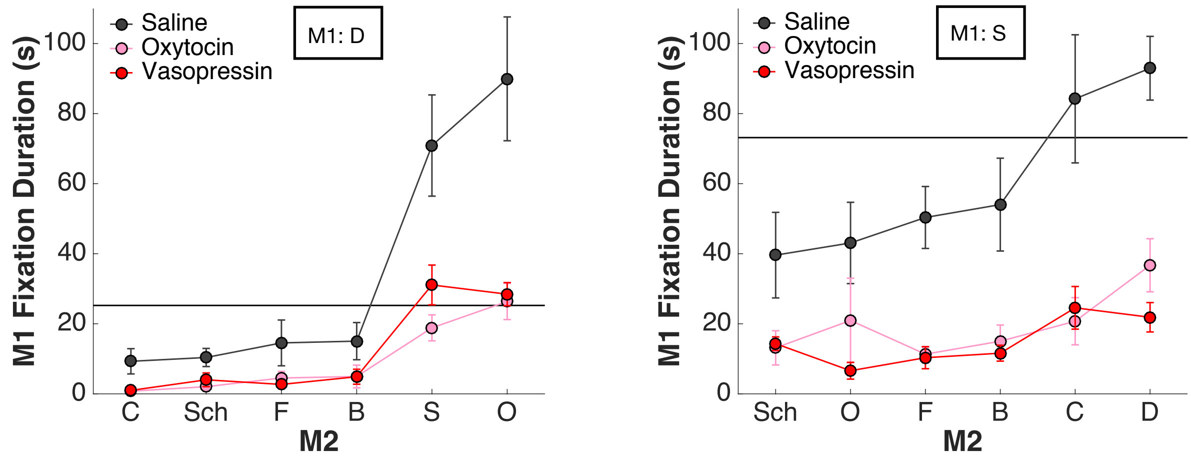

Figure 1. For the same M1 (left: D; right: S), compared with saline, OT and AVP inhalations reduce overall staring time by M1. Individuals exhibit difference in the degree of such reduction in staring. Black horizontal line: average time M1 spent staring at an empty chair; error bars: mean ± SEM.

In Jiang & Platt 2018 we have identified the main source of this individual variability to be dominance (or rather, the relative dominance between the pair of monkeys). More specifically, OT specifically reduces staring only in the most dominant M1s and M2s (for example see Rebuttal Figure 2). Even though we agree with the reviewer this nuance is very important for understanding the effects of OT and AVP administration, we do not consider this to be the most 'representative' result (per journal standard) that highlights the strength of the paradigm. We have, however, further addressed this issue in the now revised Discussion.

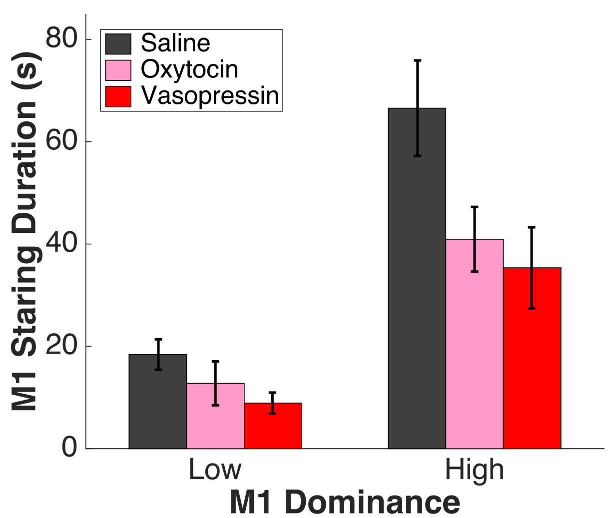

Figure 2. OT and AVP reduce staring by M1s with high dominance but not low dominance (low dominance M1, saline = 18.40 ± 2.98 s; OT = 12.78 ± 4.29 s; AVP = 8.93 ± 2.04 s; high dominance M1, saline = 66.56 ± 9.33 s; OT = 40.94 ± 6.32 s; AVP = 35.36 ± 7.91 s). Error bars: mean ± SEM.

The addition of AVP emphasizes the issue related to receptor binding/ affinity, and overall uncertainty of the mechanism that may account for the purported effects of OT on social behavior.

We completely agree with the reviewer. In fact, as mentioned in Discussion, because "…intranasal delivery of aerosolized AVP reproduces the effects of OT with greater efficacy, and that all behavioral effects are replicated when OT or AVP is injected focally into the ACCg, and "most cortical areas in primates including the ACCg lacks OT receptors but is rich in AVP receptors", we think together our results strongly indicate that "exogenous OT may shape social behavior partially through nonspecific binding with AVP receptors".

The data in figure 2 are not presented in a clear way. Although the slope estimates for the hierarchy that emerges from the analysis of pairwise staring is interesting and the plots are visually appealing, the conclusions that this is a reduction or hierarchy should be supported by statistical analysis. Overall, the ANOVA's appear to show effects but the plots are not doing justice to these data. In general, it is unclear why some of the data (that are less important) ended up illustrated in figures and other points are buried in the text.

We apologize for the confusion, but we think the reviewer may have been given an earlier version of the current manuscript. We have since significantly condensed the figures to show only the most representative results. In our opinion, both staring duration distributions (Figures 2A-B) and staring difference curves (Figures 2C-F) are equally important for understanding the main effects of neuropeptide treatment. Specifically, staring duration distributions show that neuropeptides reduced overall staring, whereas staring difference curves show that this reduction leads to a flattening of the preexisting hierarchy at both ends (i.e. the most subordinate become less subordinate whereas the most dominant become less dominant). We again refer the reviewer to our published paper for more statistical details.

The third section in the results reports a failure to replicate the effect reported in the second paragraph. OT treatment in M1 did not reduce staring in M2 yet in the previous colony (reported in the paragraph above) it did (data shown Fig 2B). This is important to address head-on and report to further emphasize that OT by itself causes little or no reliable effects. (As a matter of scientific writing, "insignificant reduction" is not a reduction at all).

We apologize for the confusion, but again we think the reviewer may be referring to an earlier manuscript, as this result is not described in the more current, condensed version of the manuscript. Nonetheless, we agree with the reviewer that there is intrinsic noise in such behavior measurements. We also think that our sample size is still relatively small, and that there may be subtle differences between the dominance hierarchy structures in different groups of subjects, which may in turn lead to different neuropeptide effects. We do want to emphasize however that in Jiang & Platt 2018, Figures 1-3 and Figures 4 were collected not only from different colonies but also in *different physical locations (the former at Duke University whereas the latter at University of Pennsylvania). We feel the fact that we could replicate most of the behavioral effects of OT inhalation in two separate colonies speaks to the validity and reliability of our behavioral paradigm.*

Figure 1A is unclear and it needs to be re-designed to better illustrate the enrollment of monkeys in OT and AVP inhalation/injection. The figure should be closer to the video #1 that shows clearly the two social partners in profile looking at or away from each other.

The reviewer's point is well understood. In the revised Figures we have chosen to keep the top-down view to illustrate the camera position (and made this point explicit in the Figure and corresponding legend), but have also inserted a profile view to better illustrate the relative positions between monkeys.

Figure 2: A and B should be labeled as inhalation and not injection trials. Why are a different number of pairs in Figure 2 C and D and E? Why aren't the slopes shown and compared?

*We apologize for the confusion, but again we think somehow the reviewer may be referring to an earlier manuscript. Regardless, the current Figures 2A-B and 2C-D do have different numbers of data points, because in Figures 2A-B each data point corresponds to one session, whereas in Figures 2C-D each data point corresponds to one unique monkey pair (n = 7*6 = 42 pairs). Figure 2E and Figure 2F also use one point to represent one monkey pair, but they have different numbers of data points from Figures 2C-D because the sample sizes are slightly different (as the first OT experiment has 7 M1s, the OT/AVP inhalation experiment has 3 M1s, whereas the OT/AVP injection experiment has 2 M1s). We refer the reviewer to our previous publication (Jiang & Platt 2018) for further statistical details.*

The ACC is a large area with multiple subdivisions that have been shown to carry our dissociable functions. Given the claim that injections of OT and AVP into the ACC caused a significant reduction in staring, the precise location of the injection would aid the interpretation of these results.

This too is explained in further detail in our previous publication (Jiang & Platt 2018). Briefly, OT or AVP were injected focally into the anterior cingulate gyrus (ACCg, not dorsal ACC or subgenual ACC), a part of ACC specifically implicated in empathy, social learning, and computations of 'other-oriented' information. We refer the reviewer to a recent review for more details on the anatomical connections and functional neurobiology of this area (Apps et al. 2016)

Finally, the treatment of this subject ignores or omits to cite several recent papers that are relevant for the ideas and the data presented here.

The reviewer's point is well understood. We have now significantly rewritten our Discussion according to the reviewers' as well as the editor's suggestions. In addition, *we refer the reviewer to our published paper for a more detailed discussion of all the relevant literature.*

The soccer game
This task is as interesting as it is difficult to train. It is remarkable that the authors managed to train two monkeys to interactively deliver and defend penalty kicks and replicate the "rules" of the game that lead to winning. The data presented in Figure 3 are convincing - the monkeys indeed discovered these rules and the cognitive process involved lend themselves perfectly for electrophysiological scrutiny. Indeed, the recordings from STS confirm that the outcome of neural recordings are only as good as the ongoing task. I think the most interesting and valuable contribution of this paper to the literature is shown in Figure 4A. The dynamic modulation of the firing rates in the course of the task in each player is worth more elaboration.

We thank the reviewer for their general positive feedback on our behavioral paradigm and its theoretical impact.

The data, however, are not presented in the best form possible. Overall the manuscript is hard to follow. For example, in Figure 3A, the cyan and blue and red and pink lines are hard to distinguish and the markers of "o" and "x" are almost invisible. This figure does not provide a good illustration of the otherwise ingenious task. Perhaps fewer trials with thicker lines and larger lettering would better serve these data.

We appreciate this critic on the graphics. This issue has been fixed per reviewer's suggestion.

Likewise, the color dots in figure 3B are not differentiable for pair 1 and 2. If the data from the two pairs are merged, the legend should not contain 2 blue and 2 pink dots; if however, the authors intended to illustrate any differences between pair 1 and 2 the colors should be clearly distinguishable.

While we understand the reviewer's point, it is somewhat a convention in awake monkey literature to distinguish task type or cell type by color while representing monkey identity using different shapes. The purpose of this practice is exactly what the reviewer has suggested, to better differentiate task/cell type, as visually grouping items by color is much easier than by shape. Most studies however will still choose to keep the monkey identity information, because some reviewers/readers will wonder how consistent certain behaviors/neural patterns are across two monkeys.

Fig 3C adequately illustrates that the kicker redirection leads more often to winning but the statistical comparison is missing. If the error bars are SEM (which should be explicitly stated in the results or figure caption) then these differences should be significantly different. The table in Fig 3D, 3E, 3F, which is an analysis of soccer strategy is irrelevant and difficult to follow.

It is stated in the corresponding figure legend that error bars in Figure 3C represents to SEM. The corresponding statistic report has been added to the main text. We have also replaced Figures 3E-F with some learning curves that we think better illustrate the dynamic and evolving nature of this game. We think Figure 3D is important and relatively straight-forward, as it clearly demonstrates that monkeys use a variety of movement combinations that are not chosen at random, but rather sensitive to the winning probability distributions, and that monkeys behave in this virtual soccer game the same way human players behave in real life soccer matches.

Figure 4. In the first plot (Fig 4A) the caption says that the gray lines indicate the beginning and end of a trial. It is unclear how these trials were aligned. Were the trials of the same duration? If yes, then the statement under 2.1.6 (general setup) is misleading and it should be stated that the trials always ended at the same time.

Yes, all trials of game play are the same duration as the X-axis speed of the ball is constant. As this is a two-player game, one of the two juicers will always go off to deliver juice reward to one of the two monkeys. As a result, the reward delivery period is always the same duration as well. We have further clarified this point in the Protocol section. Furthermore, the first two gray lines on Figure 4A mark the beginning and end of game play; reward delivery is immediately after game play (also see our response to the next question).

Fig 4 A shows very interesting data but it is unclear what happens at each time point after the second gray line that indicates the end of the trial. When does the monkey receive reward for playing the game? Are the convergent winning and losing firing rates related to reward or to having just won or lost?

The reviewer raises a good point. We have now modified this figure to better illustrate the reward period. Note that reward delivery starts as soon as one bout of play is finished, and it extends for 0.8-1.0 s (with small jittering across trials) as indicated by the third thick grey line. To answer the reviewer's question, we found most mSTS neurons do not signal trial outcome ahead of reward delivery. Most mSTS neurons start differentiate trial outcome during reward delivery and this difference in activity persists after reward delivery.

Fig 4 B The distribution of firing rates as a function of redirections for either kicker or goalie should be statistically compared. Are these significant across the population, are the different in the same neurons across redirection or lack thereof? It seems that the analyses of these interesting and unique data are unfinished. The organization of figure 4 would benefit from a different reordering (A, C, D, and B)

We have now added more statistics regarding the firing rate distributions. We have also built general linear models (GLMs) to further explore what task parameters drive mSTS firing rates. In short, as described in the manuscript, we found mSTS neurons primarily signal strategy during the task phase and trial outcome during the reward phase. In *addition, mSTS neurons are also modulated by other game-related factors such as arousal and eye gaze (see Rebuttal Figures 3).*

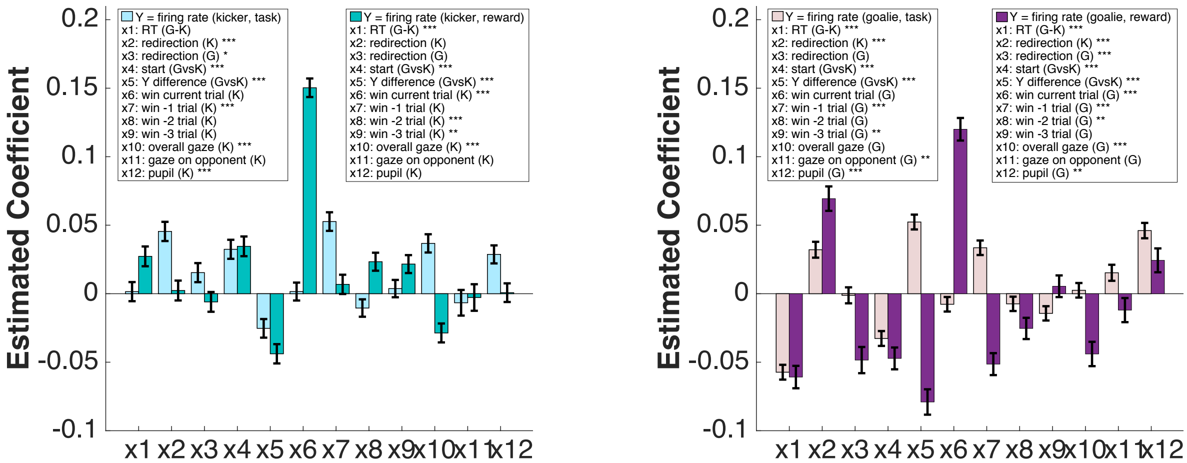

Figure 3. GLMs predicting firing rates on current trial for kicker (left) and goalie (right) neurons during live competition. Light blue, kicker task-epoch firing rate; turquoise, reward-epoch firing rate; violet, goalie task-epoch firing rate; purple, goalie reward-epoch firing rate. Error bars, estimated coefficient ± SEM.

While we have done very extensive analyses to further understand the nature of neural representation in mSTS during this game, we think the main conclusion of the current manuscript, i.e. the 'penalty kick' paradigm is a powerful, data-rich investigative tool to study dynamic social interactions, does not rely on the specific analyses performed on our electrophysiological dataset. Finally, we think Figure 4B should lead Figures C-D, as the former is about mSTS activity during task-performance, whereas the latter is about mSTS firing rate during and post reward delivery.

Overall, more detailed analyses (including more stats), more judicious decisions of what to include and not include in figures, a better anchoring of the data in the extant literature, and clearer explanation of the methods would greatly benefit both otherwise interesting and valuable studies.

We thank the review again for expressing general enthusiasm towards our paradigm. We hope the revised manuscript with new figures and additional statistic reports will address all of the remaining concerns of the reviewer.

Reviewer #2:

Manuscript Summary:
This manuscript reviews currents studies from the Platt lab that utilize interactions between monkeys to examine the social decision-making. Overall, I think this is a fair

review of their studies, but have concerns about both their broader framing of the work and some of the details of the actual studies, particularly the second experiment.

Major Concerns:
1. In terms of the 'Theoretical' framework, the authors contend that their behavioral paradigms represent natural approximations of social interactions in macaques that can be leveraged to understand the underlying neural mechanisms. This viewpoint, however, belies a misrepresentation of both the species and the paradigms being used. There is nothing 'natural' about sticking a monkey in a chair. It is a contrived setting that in turn reflects little in how monkeys would actually behave. This is not to mean that these approaches cannot inform some aspects of social cognition, but to call it natural is inaccurate. Indeed this lab has made significant contributions to our understanding economic decision making utilizing similar paradigms. Although Experiment 1 did not involve training, it is nonetheless not 'natural' to have two monkeys fixed in close proximity unable to change their location. Rhesus macaques avoid looking at each other under natural conditions, but this is the key behavioral measure for the study. Since this paper has already been published, I will not comment on the experimental details, but just make the point that it should not be characterized as reflecting 'natural' social behaviors.

We appreciate this critic from the reviewer. We have now significantly expanded our Discussion on this caveat. While it is true that all laboratory experiments, in humans or animals alike, are 'not natural' in comparison to real life, the degree to which a behavioral paradigm carries ethological validity is still a relevant question. For example, in equally controlled laboratory settings, asking a monkey to to choose between a square or a circle on the screen to indicate whether to cooperate with or betray an imaginary nonspecific is not as ethnologically meaningful as asking a monkey to choose to make eye contact with, or donate reward to, one of the other two monkeys sitting in the same room. We feel strongly that just because monkeys are seated in primate chairs (to facilitate eye movement recording, pupilometry, and electrophysiological recording), this doesn't mean one should abandon all hope and stop trying to improve upon the ethological validity of one's social paradigm. This being said, our lab also has great interest in measuring free-ranging primate behaviors in the laboratory as well as in the field. The reviewer is absolutely correct that, in comparison with naturally occurring behaviors, the best designed laboratory paradigm will still fall short in terms of flexibility and spontaneity. It is however a tradeoff one has to make from time to time, as a more restrained paradigm can afford us the opportunity to investigate more specific scientific hypotheses.

The reviewer raises a good point regarding the face-pff paradigm specifically. In our experimental setup, although the monkeys could not retreat from each other, they could look away or turn around to avoid direct confrontation. We think this relatively unconstrained but still well controlled setup affords us most of the flexibility of natural social interactions without risking actual physical contact between monkeys (for reference, see Dal Monte et al. 2016 for a very similar but more restrained setup). We have now further emphasized the caveats of this design in the revised Discussion. It is *also worth noting that, in a separate dataset (OT/AVP inhalation in females), we do have a small set of video recordings of pair-housed monkeys freely interacting with each other in the home cage in addition to our in-chair 'face-off' videos. Briefly, the free-interacting videos confirm our observation in the laboratory that neuropeptide treatment in females promotes social vigilance and aggression. While it is true that unconstrained interactions are more flexible and multi-dimensional (for example, in these videos we see monkeys approaching each other, running away from each other; or offering, accepting, or rejecting grooming), but we feel our experimental design preserves sufficient details of natural social interactions, and a slightly more unconstrained design is unlikely to alter the main conclusions of our paper.*

Finally, from a social neuroscience perspective, there is also something to be said that the closer a behavioral paradigm resembles components of real life, the more likely it is for subjects to tap into existing brain networks to solve the problem. For example, the human brain is not evolved specifically to send text messages, listen to podcasts, or play computer games remotely with other people. Yet all these activities are considered social in some way and can tap into the social brain network. Similarly, we are not suggesting monkeys play soccer with each other in the wild; instead we are arguing that our soccer task possesses enough elementary components of natural social interactions that are recognizable to monkeys that it enables us to examine the neurobiology underlying strategic social competition.

2. Since first experiment has already been published, it is reasonable that many of the key experimental details are omitted. As far as I can tell, however, Experiment 2 has not been published previously. If it has been published, the authors need to make this more obvious. Certainly this behavioral paradigm seems very intriguing. But there are simply too few details about the paradigm to accurately judge its merits. The authors present only cursory behavioral details and even less detail on the neural recordings. Without more thorough details, there is simply too little information about the study to effectively review its merits.

The reviewer's point is well understood. Review #1 has expressed similar concerns. Based on both reviewers' suggestions, we have now replaced some figures to better illustrate the dynamics nature of our task, and beefed up the Result section to include more statistics and descriptions of various aspects of the task. We did not, however, focus on expanding our neurophysiology results. While we have done very extensive analyses on this, we think the main conclusion of the current manuscript, i.e. the 'penalty kick' paradigm is a powerful, data-rich investigative tool to study dynamic social interactions, does not rely on the specific analyses performed on our electrophysiological dataset to test the hypothesis that the primate STS encodes non-perceptual social information.

Minor Concerns:
Given its relevance, I am puzzled that the Haroush & Williams study is not even cited.

We are appreciative that the reviewer has reminded us of this very reverent reference. This has been added to the revised Discussion.