

	Reviewer	Author response
	Recommender (Veli-Matti Karhulahti)	<p>We have considered all comments and completely revised the manuscript. Based on your helpful ideas, we came to the conclusion that we move the question away from replication. Instead, we have given a lot of thought to the contribution of our work to psychological issues. All comments have helped us to go deeper into the theoretical considerations. We hope that the revised manuscript is more focused and easier to follow and evaluate than the previous version. We are very much looking forward to any input you and the reviewers may have.</p>
E1	<p>1. The “scientific goal” of the study is unclear. ...what is the scientific question that the study wants to figure out. Do you wish to contribute to the theory behind Dunbar’s number? Do you wish to learn more about Shakespeare, drama, or character networks in fictional narratives? the extended replication will surely yield new useful information, but it is not clear what that means. If the original study replicates or not, what can we deduce from that, theory- or otherwise?</p>	<p>We have now fully reworked our manuscript. We now state the goals of our project more clearly</p> <ol style="list-style-type: none"> <li>1. Our study harnesses the idea of using human-made cultural products, such as literary works, to learn about the human psyche. In this sense, this generalization to the full set of Shakespeare’s plays is but a first step in that direction. We accordingly now not only focus on Shakespeare, but extend parts of the analyses to a large corpus of European drama. We hope to inspire other researchers to follow our path and to thereby contribute to a cumulative scientific process.</li> <li>2. Regarding your specific questions: We now agree that our initial analyses and results would not have contributed to the theory of Dunbar’s number. We believe that our new focus on complexity allows for a clearer link to Dunbar’s number, with the outcome of especially Study 3 either providing some degree of support for the usefulness of Dunbar’s number or highlighting some limitations.</li> </ol>

		We now frame the replication of Stiller et al. (2003) as a by-product of our work which will be reported in supplementary material, and contribute to the replicability of network-analytic methods based on drama.
E2	2. Although the MS explicitly says that is not designed to test hypotheses (Bias control), there are several criteria set for different outcome interpretations and in some cases they even lead to falsifying certain theoretical positions (as the four RQs show in the end). On the other hand, this seems like very traditional hypotheses/theory testing, sometimes with clear H1/H0/undecided interpretations. It is a bit unclear how this is different and why it has been separated from hypothesis testing and/or confirmatory work? I will list more detailed examples next.	We agree that we have not been entirely consistent in our phrasing with regard to hypothesis tests. We deleted the conflicting phrasings. We now specify more clearly to which theories and hypotheses our analyses make a contribution. We hope that the new structure also delineates hypotheses from more open-ended, exploratory, or even merely descriptive analyses more clearly. Thank you for pointing this out, as it helped us to clarify the contribution and aim of the paper.
E3	3. RQ1: “The theory is that Shakespeare’s plays and the ethnographic observation of human group size come from the same distribution.” Indeed, it is clear here that we are curious about similarity, statistically. Now, taking a few steps back, why is this similarity interesting? One could say, e.g., if similar, Shakespeare’s fiction accurately simulates real human social life (Dunbar's number serving as an auxiliary hypothesis for social life), but this would be unlikely be true due to reasons pointed by reviews showing how such simulation appears to be very inaccurate if we look at details.	We decided to move away from comparisons to ethnographic studies and group sizes.
E4	One could alternatively say, as you hint on page 5,	We agree that assuming “drama is (only) effective if it mirrors

	<p>that “drama is especially effective if it mirrors reality” i.e., if similar, one of the reasons for Shakespeare’s success is that people are able to cognitively reflect on social networks, which are (on average) similar size to theirs. Again, this seems unlikely for various reasons (which we don’t need to discuss here).</p>	<p>reality” is problematic. In the current version, our assumption is that the popularity of Shakespeare’s work is diagnostic for that Shakespeare’s plays contain character networks that are representable enough to typically allow the audience to follow the plot to a sufficient degree.</p> <p>We assume that a play would be less likely to be successful if the audience was unable to cognitively follow the narrative to a satisfactory extent.</p> <p>We now clearly state that this is a key assumption that could be tested in future work. In Stiller et al. (2003) and other papers this or similar assumptions seem to be made implicitly (see especially the last paragraph of the introduction on page 5 for more detail).</p>
E5	<p>In sum, there are interesting data and analyses, but we are not fully sure what the results will tell us (beyond statistical outcomes). The same applies to RQ2: “The theory that the average conversational clique size which is between three and four people can be found in Shakespeare plays”, and RQ4 “The theory that Shakespeare plays as dramas are in an Aristotelian view reflecting reality and show similar small world-properties in their networks.” I want to be very clear that it is fully ok to register exploratory analyses, and there is no need for confirmatory tests in RRs, but <b>currently the MS is sitting between the two sides without having fully outlined the rationale</b> (how do these exploratory analyses contribute to the literature, or what does it mean if a certain position/theory is falsified).</p>	<p>We agree. The more we thought about it, the more we noticed that the interpretation was indeed problematic and unclear. Finally, we came up with a completely revised set of four studies that build up on each other.</p> <p>We made the contribution to the literature more clear now and streamlined the entire endeavor.</p>

E6	<p>4. The fourth (anonymous) reviewer is an expert in Shakespeare as well as literature in general. Because the review was not submitted via the system, I am attaching it manually at the end of this recommendation. This (the most critical) reviewer is explicitly concerned that Dunbar's number is not suitable for drama in general due to huge genre variation. If you agree and believe that this may be true, it is one of the possible hypotheses to test and, if corroborated, it could make a major contribution to the literature on fictional social networks and their analyses.</p>	<p>We now directly address questions related to the adequacy of the number of nodes as a proxy for representability of social networks. We also make explicit that our research does not contribute to literature research (or Shakespeare research). Instead, we focus on the contribution of cultural product analysis for questions about the human psyche. The mentioning of the genre was helpful and we now describe how we attend to this in Study 3, but this is not the focus of our research.</p>
E7	<p>5. If you follow the reviewers' suggestions to set smallest effect sizes of interest, please carefully justify the SESOI by some raw effect if possible; this is a recurring matter discussed in depth at PCI RR.</p>	<p>Thank you for this comment! We have dropped the original research questions, but smallest effect sizes of interest still appear, and we make sure that we provide justifications for them.</p>
E8	<p>The reviewers also provide plenty of detailed comments on the design and methodology. Please consider them all carefully. I hope you find them useful and valuable in your revisions. Last, I want to stress that the value of this study, to me, is generally sufficient to be carried out even without the theoretical, pragmatic, or other contributions which most of my comments above address. I can see it can be a useful <b>methodological exercise</b> and resource for future scholars to learn from. However, I do hope you consider the above notes because</p>	<p>We thank you and the reviewers for all the effort and the helpful comments. It was more than medium effort, but we now hope that our manuscript contributes something. We believe that our changes to the manuscript increased the theoretical value.</p>

	with a medium effort, much more value could be generated.	
	Reviewer 1 James Stiller	
R1a	In brief this was an excellent and thorough proposed plan of research. The proposed research is interesting, and it would be good to see how the findings of Stiller et al (2003) fair across all plays. I particularly welcome revisiting the network analysis measures and distribution models. It would be interesting to hear how the researchers envisage this research <b>being extended beyond Shakespeare</b> . The following comments are in no way a negative criticism of the research proposal but instead highlighting areas where there is potential for more clarity and perhaps opportunity for a stronger connection with the source material.	Thank you very much for your kind words! As the extension beyond Shakespeare was also suggested to us by other researchers and as we could find a neatly compiled database ( <a href="#">DraCor – Drama Corpora Project</a> ) for this endeavor, we broadened the scope of our project now. With Shakespeare and the reproduction approach we delve deeply into one author and replicate your interesting findings, and by providing a summary of more than 3000 plays we show the broad picture of complexity across authors and periods.
R1b	1) Overall, the proposal sounds like a thorough critical re-test of Stiller et al (2003), however, <b>beyond critiquing Stiller et al (2003) it is not clear what the extension and application of this research is</b> . Is it about providing a 'tidier' research approach, finding out more about Shakespeare (as a genre) or furthering our knowledge about studying cultural phenomena?	In line with the comments by other reviewers, we have now reworked the contribution of our work.  Whereas we are critical of some specific aspects of Stiller et al. (2003), we see value and potential in their idea to use cultural products to infer about the human psyche and follow the suggested approach.

R1c	<p>2) <i>Slices</i>: This could do with more justification as a rationale; There is need for a more easily replicable approach, but this does have significant issues. In table 1 it is unclear what options a and b are in the associated text and therefore <b>slightly confusing on what they are trying to communicate</b>. I may have misinterpreted but they have equated scenes as the same exits and entrances (which are not). A more automated approach would remove selection error; however, the researchers from my interpretation have perhaps provided an over-simplification of the 'exits and entrances' as slices, <b>suggesting a scene is the same</b>.</p>	<p>We agree that our description was a bit confusing. We changed Table 1 on page 11 accordingly.</p> <p>In addition, we noticed that we did not clearly describe the forming of slices based on scenes. What we meant was that if we slice the text by exits and entrances, then this automatically includes the scenes, as all characters exit at the end of a scene. We clarified this on page 10.</p>
R1d	<p>Across a scene you can get an idea of clique size, i.e. who a character can 'potentially' interact with and this would be easier in an automated analysis, however, you lose the dynamic information of the scene. A scene does not reflect the dynamics of the actual scene.</p> <p>Historically, the nature of Shakespeare's plays were that they would have limited cast playing multiple characters, and within one scene you could quite frequently get small part characters that only interact with one or two characters and leave before others join the scene and never meet more than one character (e.g. messengers) – this is a common plot</p>	<p>You raise an important point, especially regarding our analysis on the dracor data (see Study 2-3), in which the text corpora were only sliced by scenes. We will discuss this as a limitation in the discussion section at Stage 2 of this project.</p> <p>We agree that a division based on scenes has problems. Nonetheless, scenes are a natural segmentation of theater plays as they often imply a change of place and time. And a segmentation by scene is encountered frequently enough (e.g. in the dracor corpus) to warrant consideration. By applying the analytical variants in study 4 we are able to provide a first estimate of the impact of this choice.</p> <p>Using multiple editions could indeed be interesting. However, we will not consider this variation in the current project just due to</p>

device to make sure the audience can track the intentions of different characters and understand that some characters, even in the same scene, might not have the same knowledge. This was a big issue in the Stiller et al analysis for A Midsummers Night Dream, where characters in theory are on stage a lot of the time but **actually asleep and not interacting**. The judgement was made to treat this as an “exit” as from a cognitive perspective for tracking drama they were not actively part of the scene. Therefore, there needs to be a clearer rationale on their term of “slice” and why a scene would be sufficient to capture such detail.

The proposed method will not pick up on these finer points and they are important as many characters are not that well connected and there tends to be only a few key characters per play (see Stiller and Hudson paper). The use of the scene as the slice misses the detail of the on-stage groups and for some plays that can result in exceptionally high connectivity that does not exist. I would therefore disagree that the scene is equivalent to the entrances and exits as this is not looking at the cognitive complexity of the plays and raises the question of what the goal of the research is. However, I am intrigued to see what they find out as the entrances and exits can vary from different editions of the plays and is undeniably an issue of interpretation (Stiller et al, did this based on **having more than one folio/ edition of each play**

feasibility and to keep the overall project more streamlined than, e.g., our first version of the Stage 1 manuscript. There are also many other factors that we cannot consider (actors playing multiple roles, text version, text deletions, ...). We will acknowledge these factors in the discussion.

	analysed).	
R1e	<p>3) <i>Analysis</i>: The analysis sounds interesting, and it would be good to see the network analysis in a more robust automated way, the original paper was largely calculated by hand as access to that network analysis software or free statistical/ modelling software was not available. The addition of Jaccard index and additional small world/ network analyses. IThe Latent Class Analysis sounds like a very good way for evaluating the distributions, however, if they are looking at doing a robust look at the similarities to naturally occurring networks/ groups it would be useful to have a clear overview of the data sources.</p>	<p>Thanks for the positive evaluation of our method. We agree that an overview and theoretically justified selection of the data sources for naturally occurring networks would have been necessary. We now reframed our contribution and made extensive changes. Thus, we do not include numbers about naturally occurring networks.</p>
R1f	<p>4) I would recommend avoid claiming Stiller et al (2003) stated facts (“As a criterion, Stiller et al. (2003) used the fact that “all the naturally occurring observations fall with the range of the ten plays, and within two standard deviations of their mean” (p. 400).”) this misrepresents the original researchers intentions. The paper was not fact but interpretation.</p>	<p>Thank you for pointing this out to us! We definitely do not intend to misrepresent the original researchers’ intentions. We dropped this sentence.</p>



R1g	<p>5) <i>Clarify how the proposed research and replication will be of use to understanding the plays:</i></p> <p>As the researchers touch on Dunbar's number etc, if this is part of what is being evaluated then there needs to be a bit more contextualisation. One of the points the researchers make about the cognitive load of the plays could be contextualised more. Stiller and Dunbar were <b>interested initially in tracking intentional stances (who knows what about someone)</b>, this can be complex and research such as Stiller and Dunbar (2006) have shown that people can struggle to do this in complex scenarios where multiple intentional stances need to be tracked. <b>The structure of the play is not the only reason for the success of the plays, but it could play a part in making the unfamiliar appear more familiar.</b> The small world network of the Shakespeare plays could provide a way to navigate this as the key characters, those that are most connected and act as "weak links" between scenes can be followed more confidently than less connected characters (obviously some poorly connected characters are essential to plot e.g. the messenger in Romeo and Juliet). Subsequent research by Stiller on the plays of <b>Agatha Christie</b> (where often characters have 100% connectivity) shows that by going away from a small world structure can <b>allow for complex story telling e.g. in detective work and tracking</b> complex intentions as this makes following viewpoints more</p>	<p>As we have made extensive changes, the role and function of Dunbar's number in our project should now be more clear (see especially Study 3). We still would like to show what cultural products can tell us about the human psyche. Since the plays are both the product of a human mind and attended to and enjoyed by human minds, inferences about the human mind become possible.</p> <p>We now clarify the goals of our study on page 5 (see also our response to E1).</p>
-----	---	---

	cognitively demanding.	
	Reviewer 2 ( <i>Matúš Adamkovič</i> )	
R2a	<p>Dear Authors, Thank you for submitting this interesting Stage 1 protocol. The RR builds upon Stiller et al. (2003) and aims to investigate whether the structure of social networks in all of Shakespeare's plays mimics real-world social networks. Before providing my comments, I'd like to disclose that I'm a psychology researcher focused on the methodology of behavioral research. I've experience applying the network approach to psychological phenomena, which is not necessarily the same as social network analysis, Additionally, I possess an average non-native English-speaker level of knowledge about Shakespeare's plays. The review will thus be based on my overall impressions of this RR and questions I had while reading it.</p>	<p>Thank you very much for taking the time to review our manuscript. We found your review very helpful.</p>
R2b	<p><b>Open science practices</b> I'd like to start by praising the authors for their level of transparency and adherence to open science practices. The publicly available materials could indeed be a valuable resource for other researchers pursuing similar research questions.</p>	<p>We are glad that you see it this way! We really hope that sharing the materials and showing how cultural products such as plays can be analyzed in a transparent and reproducible way gives some researchers the courage to pursue similar ideas - and to build on our methodology.</p>
R2c	<p><b>Abstract and introduction</b> The abstract is succinct and outlines the study's rationale. Even though it's Stage 1, the abstract would benefit from specifying the research questions.</p>	<p>Thanks. We completely revised the manuscript and now suggest a set of four studies.</p>

R2d	<p>In the introduction, the authors reference the upper bound of the size of social groups. Can the authors provide a more detailed explanation of this limit, especially in the context of today's society (e.g., usage of social networks) and its heterogeneity across cultures?</p>	<p>We changed the focus of our research (see also our reply to R1e).</p> <p>We now discuss upper bounds in social networks as a more abstract concept, noting that most simple metrics such as the size of social groups are only useful metrics conditional on several assumptions (page 3-4). For example, even with a fixed group size, the overall complexity of a network and the information required to represent said network can vary based on further structural characteristics. We thus focus now on the Kolmogorov complexity of the network and test how it relates to the number of nodes (to which Dunbar's number refers). We also explore in Study 3 how heterogeneity across cultures, literary genres, and epochs etc. can manifest in different ways, with different implications for complexity.</p>
R2e	<p>Can the authors elaborate on why Shakespeare's plays (besides the ten investigated by Stiller et al.) were chosen for this analysis. What are the characteristics that make these plays unique/suitable for such a study? Furthermore, can the authors summarize (methodological) differences between the present study and the study by Stiller et al.? From a layperson's perspective, I also wonder about the rationale of examining the research questions on all 37 plays. <b>Do the authors assume that this mirroring of social interactions is universal across the plays and can the heterogeneity in the genres, contents, and complexity of character interactions be disregarded?</b> While I very much appreciate the authors' efforts to create methodologically and technically rigorous workflow, I'd like these substantial questions clarified.</p>	<p>In reaction to this and other comments, we now include a corpus of more than 3000 plays, and compare the complexity of Shakespeare's plays to it.</p> <p>The reason behind choosing Shakespeare's plays is their popularity.</p> <p>Our assumption is that the popularity of Shakespeare's work is diagnostic for that Shakespeare's plays contain character networks that are representable enough to typically allow the audience to follow the plot to a sufficient degree.</p> <p>We also assume that Shakespeare's plays differ in their complexity. We are now able to investigate the distribution and put it into perspective with regard to European theater plays.</p>

R2f	<p>A minor note: Although I mostly agree with authors' definitions of reproducibility, replicability, robustness, and generalizability, I'd suggest adding a reference (e.g., <a href="https://doi.org/10.1146/annurev-psych-020821-114157">https://doi.org/10.1146/annurev-psych-020821-114157</a> or NASEM's report <a href="https://www.nationalacademies.org/our-work/reproducibility-and-replicability-in-science">https://www.nationalacademies.org/our-work/reproducibility-and-replicability-in-science</a>).</p>	<p>Thanks for these references. As we changed the contribution, we could not incorporate them here, but they will be useful for us in future work.</p>
R2g	<p><b>Reproducibility, generalizability, and robustness testing</b>  The authors state that there are three major pathways (i.e., the choice of plays to be included in the analysis, how the play is segmented into time slices, and the criterion for tie-formation) that could determine the results. I fully agree. These researcher degrees of freedom cover the selection of plays and, in essence, data pre-processing. Based on my experience with network analysis of psychological phenomena, the resulting parameters are often sensitive to analytical choices (e.g., estimator selection, setting tuning parameter/s, etc.). If these analysis-related choices could determine the results (i.e., no single optimal network construction algorithm exists), would it be possible (and make sense) to incorporate this into the code and multiverse the results?</p>	<p>We think that the three factors (now only two) that we chose for Shakespeare's data are theoretically important. With character networks, most measures do not require tuning parameters etc. There are degrees of freedom, however, with regard to Kolmogorov complexity. We make our suggested choices transparent and are open for any input you or others may have.</p>
R2h	<p><b>Non-registered analyses</b> When comparing the number of speaking characters, the authors propose to use paired Wilcoxon tests. I'd suggest using Welch's t-tests instead, set SESOI, and conduct equivalence testing (or a Bayesian analysis) in addition to NHST. I've never seen Weber fraction</p>	<p>Thank you for this helpful comment. This analysis is not part of the current set of planned hypotheses anymore.</p>

	used outside cognitive psychology research – it looks interesting and promising.	
R2i	<p><b>Registered analyses</b></p> <p>Overall, as far as my expertise goes, I find the registered analyses technically sound. For registered analysis no. 2, I suggest reporting not only the frequency of three- or four-character configuration but extending it to a distribution of all character configurations. Perhaps a formal test (e.g., chi-square) can be a useful addition to determine if the observed frequency of three or four characters per time slice differ from what one would expect by chance. This can also be useful to examine if the distributions of characters per time slice differ across the analytic variants. For registered analysis no 3, I'd suggest the authors take a look at the NetworkComparisonTest (van Borkulo et al., 2017) package for R. The permutation-based approach introduced in the package can be helpful to answer the pursued research question in a more precise manner.</p>	Thank you for this helpful comment. This analysis is not part of the current set of planned hypotheses anymore.
R2j	<p><b>Final remarks</b></p> <p>From a technical and methodological perspective, the present RR is rigorous and can be impactful in a way that it has the potential to greatly help other researchers who pursue similar research questions. From a substantial viewpoint, I think that the RR would benefit from several clarifications. These would greatly help non-specialized readership to get a better understanding of the paper and its rationale</p>	We thank you for reviewing our stage 1 report. We have completely revised the manuscript and included many of your helpful advice. We clarified our aim and the methodology.

	<i>Reviewer 3 (Tomáš Lintner)</i>	
R3a	The authors present a research aiming to link the overlap between human's capacity to follow on social relations and its presentation in cultural artefacts. The authors explain their intentions in a clear, concise, and transparent manner. The report provides sufficient justifications for the objectives and for the methods planned to be used.	We thank you for reviewing our manuscript.
R3b	There are, however, a few comments/questions I would like to share:  - the authors largely introduce their research on the paradigm of Dunbar's number. The authors provide a short introduction to the paradigm and support it with seminal works. However, at this stage, the authors neglect a large research array of <b>research standing in contrast to the paradigm of Dunbar's number</b> . In future, it would be useful if the authors could briefly problematize the paradigm in the introduction/theory, and not just present it as granted.	We incorporated your advice to reference work debating and criticizing Dunbar's number (see page 3). We further test how the number of nodes relates to the complexity of networks across plays (see also our reply to R2d).
R3c	- on p. 6, the authors start describing the Stiller et al.'s (2003) reports on the cliquishness and the small world properties which Stiller et al. (2003) relate to "naturally observable human social network	We now changed our focus and state how focusing on the complexity of a play can yield insights about human cognitive capacity.  Presence of characters does not necessarily indicate information

	<p>properties". I find it hard to follow on why it makes sense to analyze and interpret the connectivity within the play networks in which the ties between the characters represent not their social relations, but the occurrence on the stage at the same time, in a similar manner than regular social connections between people would be analyzed and interpreted. Like the authors themselves write and rely on Latora &amp; Marchiori (2001), in real-world social networks, <b>connectedness usually denotes the ability to spread information</b>. However, the <b>presence</b> of characters at the stage at the same time creates a very <b>different type of ties</b>. It would be useful if the authors could <b>clearly define what the ties within the plays represent</b> and how the structural network properties created by these ties will be interpreted.</p>	<p>spread.  We would argue that the co-presence represents a connectedness of two roles due to being observed at the same time. This is not far away from general social network research because often ties are influenced by exactly this kind of standing together (physical proximity). Nevertheless, for the interpretation of the properties we are aware that some concepts are used for the description of the character networks that do not have the same meaning as for real social networks. But we want to highlight that the difference in meaning is more important for the implication of the results than for the general way the index is understood (see also R3e).</p>
R3d	<p>- on line 244 of the code and further, the authors plan to drop the "All" speaking characters. If I understand this correctly, this is because at these moments, all characters are speaking at the same time? If yes, I probably understand why the authors would want to drop all those occurrences, but I think it would be useful to explicitly describe that, because if I understand that correctly, it can influence the results of the analysis.</p>	<p>This is correct. We also agree that this can have an impact on the network characteristics.   We describe it explicitly in the manuscript now (page 12).</p>

R3e	Overall, the report is comprehensive and the R code makes sense. My main suggestions for improvements revolve around the future interpretation of data since the authors will be dealing with a very different of type of ties and the frequently-used structural indices may have different meaning compared to, for example, interpersonal relationships.	Thanks also for having a look at the R code and providing us feedback to that. In general, we agree that the type of relations used here has a specific role but less in terms of a different meaning, but more in their meaning for the consequences of the results. We aim at inferring something about the cognitive capacity to follow a play.
	Reviewer 4 (Anonymous - Shakespeare expert)	
R4a	Let me start by saying that I was asked to report on this research as a literary expert. I will thus not discuss the stastical side of the authors' work, but only their potential interest and validity for literary analysis.	Thank you very much for agreeing to review our manuscript.
R4b	1A. The scientific validity of the research question(s)  The study they intend to replicate had virtually no value for literary study. The selection of the 10 plays <b>made little literary sense</b> ,	Thank you for sharing your perspective as a literary expert. We agree. We make clear that we do not approach human cultural products as literature researchers, but from a psychological perspective.
R4c	because a) it focused only on those plays that most <b>coveniently agreed with the "Dunbar number"</b> ,	We now consider many more plays and relate the number of nodes in a network to the complexity of a network.
R4d	and b) ignored what is crucial from the viewpoint of drama (and of dramatic networks), <b>namely the difference in genre</b> . On a) this replication may indeed prove useful, in testing, and almost certainly falsifying, the original study. On b) no, because the	We thank you for this comment and explore in Study 3 the heterogeneity across genres, and other variables with different implications for complexity (See R2d).



	<p>study shows to have a <b>total disregard for dramatic genre. (Genre is meaningful because comedies have always a much higher density than tragedies which have a much higher density than histories; ignoring this initial fact creates only confusion.)</b></p>	
R4e	<p>In addition, the (infrequent) moments in which the study mentions literature its categories – and references – can only be <b>described as primitive</b>; even when they refer to quantitative and/or network analysis of drama they mention very peripheral studies, and ignore crucial ones – such as Yarkho’s on speech groups.</p>	<p>Thank you for your perspective and for pointing us to Yarkho’s work. We found it very interesting, but do not think that the revised version of our Stage 1 manuscript faces the same challenges as the previous version.</p>
R4f	<p>1B. The logic, rationale, and plausibility of the proposed hypotheses (where a submission proposes hypotheses)</p> <p>I do not believe a hypothesis is being proposed.</p>	<p>We now are mainly exploratory and do test only one specific statistical hypothesis in Study 3.</p>
R4g	<p>1C. The soundness and feasibility of the methodology and analysis pipeline (including statistical power analysis or alternative sampling plans where applicable)</p> <p>I am not qualified to evaluate that.</p>	<p>Thank you for your helpful input on the issues that fall under your expertise. This kind of input is very valuable to us.</p>
R4h	<p>1D. Whether the clarity and degree of</p>	<p>Following this and other comments we have completely revised</p>

	<p>methodological detail is sufficient to closely replicate the proposed study procedures and analysis pipeline and to prevent undisclosed flexibility in the procedures and analyses</p> <p>It might ; I am not qualified to judge. But the question assumes that the original study deserves to be replicated – an assumption I personally consider groundless.</p>	<p>the aim of our work. We discuss the potential merits of the approach described in the original study for fields such as psychology (see page 3).</p>
R4i	<p>1E. Whether the authors have considered sufficient outcome-neutral conditions (e.g. absence of floor or ceiling effects; positive controls; other quality checks) for ensuring that the obtained results are able to test the stated hypotheses or answer the stated research question(s).</p> <p>I am not qualified to evaluate that.</p>	<p>We would like to thank you for sharing your perspective with us!</p>