Cell Systems

Transparent Peer Review Record

Emergence of synchronized multicellular mechanosensing from spatiotemporal integration of heterogeneous single-cell information transfer

Amos Zamir, Guanyu Li, Katelyn Chase, Robert Moskovitch, Bo Sun, Assaf Zaritsky

Summary

Initial Submission: Received Oct 2, 2022
Preprint: doi.org/10.1101/2020.09.28.316240
Scientific editor: Bernadett Gaal, DPhil

First round of review: Number of reviewers: 3
3 confidential, 0 signed
Revision invited Dec 22, 2020
Major changes anticipated
Revision received Jan 14, 2022

Second round of review: Number of reviewers: 2
2 original, 0 new
2 confidential, 0 signed
Accepted Jul 7, 2022

Editorial decision letter with reviewers' comments, first round of review

Dear Assaf,

I hope you are well! I'm enclosing the comments that reviewers made on your paper, which I hope you will find useful and constructive. As you'll see, they express interest in the study, but they also have a number of criticisms and suggestions. Based on these comments, it seems premature to proceed with

the paper in its current form; however, if it's possible to address the concerns raised with additional experiments and/or analysis, we'd be interested in considering a revised version of the manuscript.

As a matter of principle, I usually only invite a revision when I'm reasonably certain that the authors' work will align with the reviewers' concerns and produce a publishable manuscript. In the case of this manuscript, the reviewers and I have a few make-or-break concerns.

1. The reviewers raise possible alternative interpretations, which we see as crucial to address. These include:

   Reviewer 2's major point 1, raising the possibility that synchronisation may not be due to cell communication;

   Reviewer 3's major point about how synchronisation is inferred (the section with the header 'Asynchronization');

   Reviewer 3's major point 3 about whether the cells had enough time to sufficiently recover between cycled to make the assumption of independence between cycles valid and to reasonably make your conclusions with regards to memory.

2. The reviewers request clarification of and justification for the choice of metrics and assumptions as well as choices of normalisation and representation of the data. These concerns and questions are critical to address and clarify in the next revision to ensure that the reviewers are satisfied that there are no methodological or conceptual flaws in the paper and that the claims are absolutely clear to the reviewers.

While it will also be important to ensure that the paper is more accessible to a broader audience, and I would encourage you to have a go, I can work on this with you at a later stage.

In addition to the concerns I've detailed above, I've highlighted portions of the reviews that strike me as particularly critical. I'd also like to be explicitly clear about an almost philosophical stance that we take at Cell Systems…

We believe that data are of primary importance; text, story, and conclusions are secondary. When forced to make a choice, we would rather have text that hews closely to observations and reflects data precisely than a gloss that obscures alternate interpretations. Please keep this in mind when addressing the points the reviewers raise about choices of normalization and data presentation.

As you address these concerns, it's important that you and I stay on the same page. I'm always happy to talk, either over email or on Zoom, if you'd like feedback about whether your efforts are moving the manuscript in a productive direction. Do note that we generally consider papers through only one major round of revision, so the revised manuscript would be either accepted or rejected based on the next round of comments we receive from the reviewers. If you have any questions or concerns, please let me know. More technical information and advice about resubmission can be found below my signature. Please read it carefully, as it can save substantial time and effort later.

I look forward to seeing your revised manuscript.

All the best,

Bernadett

Bernadett Gaal, DPhil
Scientific Editor, Cell Systems

Reviewers' comments:

Reviewer #1: The authors analyze mechanosensing via calcium dynamics in endothelial cell monolayers, and characterize how heterogeneity with respect to cellular function and intercellular communication affect multicellular synchronization. The partitioning of the cell population into distinct classes of behavior (individual, leader, follower, etc.) provides an interesting way of understanding how single cell dynamics play into larger population level function.

The manuscript is written clearly, though there is a significant overlap between the methods and results sections. The figures are easy to understand, and communicate the results described in the text.

MAJOR REVISIONS/MAJOR COMMENTS

The introductory discussion on heterogeneity and noise needs *significant* revision and clarity. Phenotypic heterogeneity, for example, is not the same as extrinsic noise. Extrinsic noise is the variation in expression/abundance of a species across cells in a population. It may result in phenotypic heterogeneity, but the two are not the same. It would strengthen the authors' work to distinguish these points and be more precise/deliberate throughout the manuscript, especially since the work is grounded and motivated on these important concepts.

Further justification on the choice of synchronization metrics is needed. Since the current metrics are some measure of un-normalized standard deviation, the "asynchronization" measure and max-min metrics are without bounds. Classical order parameter metrics that are bound between 0 and 1 are applicable to such time-series data and would be easier to interpret. Also, since the experiments seem to be run up to 1200 seconds, the choice to use only the final 200 seconds to calculate these metrics seems odd and is not justified. In most systems, synchronization dynamics are often most "active" and relevant immediately upon application of a perturbation (rather than at the end of the experiment when the system is approaching steady state).

Additional major gaps (outlined below) need to be addressed/justified.
- Why are next to nearest neighbor cells included in calculating transmission and receiver scores?
- What might be a biological/physical motivation to go beyond a topological distance of 1?
- It seems that counting next to nearest neighbors while calculating these scores presents double counting with respect to intercellular communication. How do the authors mitigate this bias?
- How might intercellular communication be realized in this context?
- While the GC scores provide a statistical basis to suggest it exists, how might such communication happen physically/chemically? Connecting it back to mechanism--even if hypothetical--would help

ground the study.

The use of z-scores should be justified since the distribution of GC coefficients seems to be asymmetric, both for transmission and receiver scores.

What could one infer from the fact that the observed fraction of hubs is greater than P(z_tranmission >0.5)*P(z_receiver >0.5) other than the two scores being non-independent? Why might a cell capable of receiving information also be more likely to transmit information and vice versa?

In calculating the EMD metric for bulk heterogeneity, why is a uniform distribution used as the baseline? What might suggest an equal distribution of roles in the population/why might it be a useful prior?

In Fig. 2E, the authors mention that "These results suggest that increased fraction of communication hub cells along with decreased fraction of individuals were associated with an improved synchronization process". It seems reasonable that when more cells participate in communication and fewer take on an exclusively individual role, a more coherent outcome is likely. Isn't this expected?

MINOR REVISIONS, SUGGESTIONS & COMMENTS

The classification of individual/common/leader/follower/hub cells happens when the scores are either <-0.5/(-0.5,0.5)/>-0.5. A negative sign needs to be added where appropriate.

Replace "stableness property" with "stability".

A summary of how many cells (as a fraction of the whole) passed both stationarity tests in each experiment would better support the results.

Since heterogeneity can take on many forms, it would be useful to anchor the use of the term to a specific context throughout the manuscript and to use them in a better defined and deliberate manner.

What could be an explanation for the sharp increase in the hub to individual ratio in the last two cycles?

Reviewer #2: The paper by Zamir et al addresses an interesting question related to the synchronization of Ca2+ signaling among a population of endothelial cells. The questions raised are interesting and the overall approach used in the paper is appropriate. The results claim to show how synchronization initiates as local communication that propagates toward the population as a whole. If correct, these results are of interest to the broader community, and the paper a great fit for Cell Systems. However, there were a few things that confused me about this paper's methodology and analysis. These confusing points have to be addressed before I can endorse publication. If the author could better support their claims and clarify the key points raised below this paper will be a great candidate for publication at Cell Systems. However, there is a chance that the confusion is not just a misunderstanding but points to a logical flaw and if that is the case the paper should not be published. My recommendation is to return this to the authors and allow them to respond. Note that no additional experiments are requested.

Major points.
1. Synchronization might not be due to communication.
This is a tricky one and my main concern about the interpretation of the results. Let's say that cells are not communicating at all. But, each one of them, independent of the others, is all getting better at expecting a spike, i.e. change in flow rate. This could be something simple like a negative feedback loop on receptor occupancy, intracellular regulation of some kinase that regulates ATP receptors, or something like that. In that case, all cells will still look like they are synchronizing even though they are not communicating with each other. All the global measures could still look similar to figure 1 and 3BC. In that case, since cells are more coordinated, they should also be a bit more predictive of each other and I can imaging that the GC edges might still be significant even if they are not really talking to each other. They are significant because other cells (that are becoming more correlated with each other) will carry some information on the cell's future response. What I was missing is an analysis that shows that all (most) the GC-edges disappear when you do spatial permutation. Yes, I am aware of figure S8A but it shows the data in a compounded way (since it takes into account the hub/individual definition) so it's harder to interpret. In my understanding, any GC-edge that is still there after spatial permutation is not "real" communication. If there are a lot of these after such permutation, how can one interpret these edges as truly showing information flow? This is clearly shown in figure 5A where the bulk of the changes in P(GC-edge) is easily explained by the permuted data and only a tiny fraction is really left (see inset). If indeed cell cell communication is the key factor in creating synchrony, then the difference from random (spatially permuted) should be much much bigger.

2. Information flow - significance vs effect size.
Another thing that confused me is the use of statistical significance to decide if a GC-edge is there rather than the actual effect size. Two cells can communicate with each other in a way that barely influences their Ca2+ dynamics, yet the effect is not random, i.e. p-value will be below the threshold. One of the reasons I suspect that the p-values are so small is that you are treating each timepoint as independent in the way you set up the F-test. This means that the sample size is very large and a very tiny effect could still be called significant. Defining cell to cell communication based on effect size rather than p-values will be more appropriate and could be informative in explaining how much cells really influence each other. This might compound the interpretation of the GC-edges.

3. Cellular "roles" definition.
I am confused about the need to normalize the transmission/receiver score and why the normalized version is used to define the different "tasks" cells putatively take. As cell become more uniform in their response over activation cycles, the variance will shrink and the z-score will be inflated. I understand that there is a lot experiment to experiment variability and therefore the author felt some need to normalize, but I think that z-score could mess up the interpretation as the variance is not expected to stay the same across cycles. The change in the spread of points in Figure 3E could easily be related to normalization and not underlying. I'm not sure exactly how, but a version that takes into account spatial permutations for normalization seems more appropriate to me.

Minor/stylistic comments - some of these points are a matter of opinion and I try to point out which is which.

1. (opinion) Reduce the number of measures defined to focus on the ones that matter most. Some of these seem pretty arbitrary and the correlation between them (for example, sync rate vs rel var) are true almost by definition.

2. In a few places, you compared P(hub) vs P(z-transmissoin>0.5) & P(z-reciver>0.5) isn't that how a hub is defined in the first place?

3. (opinion) the abstract was missing a sentence of two that frame the question of the paper. The first few sentences go straight to describe what was done and what was found. It would be better to explain why and place this is context.

4. (opinion) in the discussion the first paragraph is a bit broad and is not sufficiently focused on summarizing the results. A lot of that is happening in 2nd paragraph instead. Some of the content of the 1st paragraph that describes the broader context ("This model-free data-driven approach can be applied to a broad set of biological systems…") should come later-on.

5. The reason behind the choice to use derivative instead of directly using Ca2+ value was not fully explained.

Reviewer #3: Review: Zamir et al.

Summary: This paper poses a really fascinating question about the extent of signaling information propagation and correlation within a responsive and heterogenous tissue. This is quite different from other studies I have seen that primarily emphasize mechanical force coupling, and it tackles the problem of heterogeneity head-on. It lays out a series of metrics and approaches to try to break down and quantify the coupling effects using approaches from statistical mechanics and information theory, and the level of quantitative analysis seems excellent and thorough. However, I confess that I found it difficult to follow at points, especially justifications for why specific assumptions were made, and for what I felt was a lack of sufficient explanation for a broad audience. I also had some specific technical concerns about the quantifications and experimental approaches that I think the authors can easily address. Overall, I think this is an exciting and important paper for the field.

My comments below track the order of the figures.
1. Figure 1 and establishment of baseline measurements
The microfluidic shear assay is well conceived at a hardware level, but the data presentation is somewhat confusing.
i. Stepped pressure vs. fixed pressure
I understand that stepped flow speeds can be generated (Fig. 1B) and that this results in changed in the calcium indicator intensity (1C), but I think the authors need to show clearer data for 1C of just 0.2 Pa in single step configuration rather than showing 0.2◊0.3. If all of the paper is conducted at 0.2 Pa, show us only 0.2 Pa held for 300 seconds and scaled so the dynamics are clearer.

ii. R'_i and standard deviation presentation (1D/E)
Where does stimulation start? Is t=0 when the 0.2Pa pressure shift is first applied? If so, I think the plots should be redone to show control data (without any shear stimulation) and then clearly indicate where the shear perturbation starts. I would show control data going back at 1 min. I know that R(t) normalizes by F_o, but I'm also confused about why R(t) starts at a non-zero value, and I think this has to do with the control period pre-stimulation not being shown.

iii. The three metrics

The attempt to carefully define synchronization metrics is critical, but I had some confusion about why these specific metrics were chosen and how the assumptions were justified (e.g. 200 sec windows, etc.). I think the main text itself needs to have a more complete discussion of the metrics and why they are useful/what assumptions were made. I also did not find the additional discussion in the Methods to clarify the metrics.

* Asynchronization:
o Why pick the last 200 seconds instead of more objectively defining steady state?
o The implication seemed to be that a low Asynchronization implied synchronization (opposite of asynch. is synch.), but it seems what it actually implies is that many cells have settled down and have minimal rates of fluorescence change, which doesn't seem to imply that those cells are necessarily coupled, just that their fluorescence intensity is no longer fluctuating (leading to low R'(t) and lower standard deviations. I am likely misunderstanding something, but it would help to have a better discussion of this in the next.

* Relative variability
o See above—not clear to me as a reader why the lower standard deviation inherently implies synchronization.
o Max standard deviation - min standard deviation seems to capture the total spread of the variation, and that seems important, but I don't understand what this is 'relative' to since it's not a dimensional quantity. Does this make more sense as a ratio, or normalized with respect to control data?

* Synchronization rate
o Why 200 seconds post-peak standard deviation? Could you instead measure the time for the system to go from peak deviation to minimum deviation?
o I did not understand the definition of the ratio of the integral over 200 s post-peak to the 'theoretical upper bound'. The figure and text were difficult to follow in this regard.
o What units does this have that make it a rate?

* Additional parameters
It seems like it would help to compare or normalize to control data. It would also help to understand if the cells ever actually truly equilibrate back to their control fluorescence intensities.
o I think it would help to show the ratio of sigma_steady_state to sigma_ctrl to clarify this. If the system equilibrates, this ratio should go to 1. This seems like a more objective metric than averaging over 200 seconds.

2. Information transfer and the 5 classification types (Fig .2)
* It would help to see a fraction plot (or pie chart) breaking down the fractions of the population that belong to each of the 5 defined categories
* Fig. 2C is a bit confusing . It's hard to visually map the two images to look at the relative ratios of transmission/receiver. Can a ratio be shown over these two to give a better sense of where the 'hubs' are and how much a given cell can be both transmitter/receiver?

3. Periodic flow and memory (Figs. 3/4)
I have a specific biological and experimental concern here, and this is my primary concern with the paper and interpretations.
* Show how the system responds to a single, 2 minute shear pulse with sufficient time to allow equilibration (plot standard deviation and R' for a single cycle with an additional 2-3 minutes after the shear is stopped). This will reveal the equilibration dynamics, as well as the general system response

for a short pulse.
* Include control data prior to stimulation: As before, go back at least 1 full period (2 min+) for the plot above and the plots already shown.
* Assumption of independence: The authors state that each cycle was treated as independent, and that each cycle was run over 2 minutes. However, when I look at the midline axis it looks to me like the R' signal never actually reaches equilibrium between cycles, especially early on. To me, this implies that the cells are still recovering from the Nth cycle when the N+1 cycle starts. While some of the cycles look like they may cross zero before the next begins, it's such a fine distinction that it still doesn't seem like the cells have had a chance to recover, meaning their behavior in the N+1 cycle may be biased by their behavior in the Nth cycle. This seems to matter because the analysis of cellular memory and how cells maintain or transition across roles depended on the assumption of independence across cycles. While the effect of tightening standard deviation is clearly real and meaningful regardless of independence, it's not clear to me how this relates to the concept of 'memory' especially with respect to questions of if a given cell has a preferred behavior. Would a cell that is a strong transmitter in cycle 1 still be such in cycle 2 if it were allowed to fully relax between cycles? Again, I may be misinterpreting this, but it seems like the experiment would need to be re-run with sufficient rest time between each cycle if independence is a necessary condition.

4. Topological distance discussion (Fig. 5)
I'd like to see more discussion of what these plots mean, what spatial permutation entails, and how to interpret the data in a broader context.

Authors' response to the reviewers' first round comments

Attached.

Editorial decision letter with reviewers' comments, second round of review

Dear Assaf,

I'm very pleased to let you know that the reviews of your revised manuscript are back. Please accept our apologies for the delay and thanks for your patience once again. We were unable to get feedback from Reviewer 1, but we are happy to move forward towards publication in light of the comments from the other reviewers. You will see that Reviewer 3 has some remaining concerns, which will need to be addressed before we can formally accept the manuscript for publication.

In addition to the final comments from the reviewers, I've made some suggestions about your manuscript within the "Editorial Notes" section, below. Please consider my editorial suggestions carefully, ask any questions of me that you need, make all warranted changes, and then upload your final files into Editorial Manager.

As you look forward to acceptance, please do consider submitting one of the protocols you've developed in this paper to STAR Protocols, or extending this offer to one of your trainees. STAR Protocols is an open-access peer-reviewed journal from Cell Press aiming to make the daily work of the scientific researcher easier by providing complete, authoritative, and consistent instructions on how to conduct experiments. The primary criteria for publication in STAR Protocols is usability and reproducibility. You can check out their most recent protocols here. If you have any questions, please email starprotocols@cell.com.

I'm looking forward to going through these last steps with you. Although we ask that our editorially-guided changes be your primary focus for the moment, you may wish to consult our FAQ (final formatting checks tab) to make the final steps to publication go more smoothly. More technical information can be found below my signature, and please let me know if you have any questions.
All the best,

Bernadett Gaal, Ph.D.
Editor-in-Chief, Cell Systems

Editorial Notes

Transparent Peer Review: Thank you for electing to make your manuscript's peer review process transparent. As part of our approach to Transparent Peer Review, we ask that you add the following sentence to the end of your abstract: "A record of this paper's Transparent Peer Review process is included in the Supplemental Information." Note that this ***doesn't*** count towards your 150 word total!

Also, if you've deposited your work on a preprint server, that's great! Please drop me a quick email with your preprint's DOI and I'll make sure it's properly credited within your Transparent Peer Review record.

Abstract:

I found your abstract a little unclear. Please give it another try. In particular, defining your model system would help anchor and orient the reader. Please also keep in mind that the Abstract should not exceed 150 words (the transparent peer review statement is not included in this count). Note that most effective abstracts have the following structure:

> *[One sentence of background.] However, X. [X essentially presents the problem that you will solve.] Here, we [solve X]. [A compound sentence that begins with a methodological phrase and ends with a phrase that describes the results that the methodology produces (e.g. "Using a combination of method Y and method Z, we show something new about X.").] [Sentences that describe key results and include carefully selected details that allow the reader to place the key results within a broader context.] [Optional: Include a sentence that "zooms out" to suggest future experiments or demonstrate impact.] [Optional: If your manuscript introduces data or software, conclude with, "[Name] is freely available for download at [DOI]." (150 words, max.)*

For a complementary view on how to write abstracts, read: http://crosstalk.cell.com/blog/how-to-hook-an-audience-with-a-great-abstract.

Please let me know if you'd like feedback on your edited Abstract before you resubmit the paper!

Manuscript Text:

- House style disallows editorializing within the text (e.g. strikingly, surprisingly, importantly, etc.), especially the Results section. These terms are a distraction and they aren't needed—your excellent observations are certainly impactful enough to stand on their own. Please remove these words and others like them. “Notably” is suitably neutral to use once or twice if absolutely necessary.

Figures and Legends:

Please look over your figures keeping the following in mind:

- When color scales are used, please define them, noting units or indicating "arbitrary units," and specify whether the scale is linear or log.
- Please ensure that every time you have used a graph, you have defined "n's" specifically and listed statistical tests within your figure legend.
- Please ensure that if you include representative images within your figures, a "representative of XXX individual cells"-type statement is made in the legend.
- Please ensure that all figures included in your point-by-point response to the reviewers' comments are present within the final version of the paper, either within the main text or within the Supplemental Information.

STAR Methods: Note that Cell Press has recently changed the way it approaches "availability" statements for the sake of ease and clarity. Please revise the first section of your STAR Methods as follows, noting that the particular examples used might not pertain to your study. Please consult the STAR Methods guidelines for additional information.

RESOURCE AVAILABILITY

Lead Contact: Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Jane Doe (janedoe@qwerty.com).

Materials Availability: This study did not generate new materials. -*OR*- Plasmids generated in this study have been deposited at [Addgene, name and catalog number]. -*OR*- etc.

Data and Code Availability:

- **Source data statement** (described below)
- **Code statement** (described below)
- Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

Data and Code Availability statements **have three parts and each part must be present. Each part should be listed as a bullet point, as indicated above.**

Instructions for section 1: Data. The statements below may be used in any number or combination, but at least one must be present. They can be edited to suit your circumstance. Please ensure that all datatypes reported in your paper are represented in section 1. For more information, please consult this list of standardized datatypes and repositories recommended by Cell Press.

- [Adjective] data have been deposited at [general-purpose repository] and are publicly available as of the date of publication. DOIs are listed in the key resources table.

- [Adjective or all] data reported in this paper will be shared by the lead contact upon request.

Instructions for section 2: Code. The statements below may be used in any number or combination, but at least one must be present. They can be edited to suit your circumstance. ***If you are using GitHub, please follow the instructions here to archive a "version of record" of your GitHub repo at Zenodo, then report the resulting DOI. Additionally, please note that the Cell Systems strongly recommends that you also include an explicit reference to any scripts you may have used throughout your analysis or to generate your figures within section 2.***

- All original code has been deposited at [repository] and is publicly available as of the date of publication. DOIs are listed in the key resources table.

Instructions for section 3. Section 3 consists of the following statement: Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

In addition,

STAR Methods follows a standardized structure. Please reorganize your experimental procedures to include these specific headings in the following order: LEAD CONTACT AND MATERIALS AVAILABILITY (including the three statements detailed above); EXPERIMENTAL MODEL AND SUBJECT DETAILS (when appropriate); METHOD DETAILS (required); QUANTIFICATION AND STATISTICAL ANALYSIS (when appropriate); ADDITIONAL RESOURCES (when appropriate). We're happy to be flexible about how each section is organized and encourage useful subheadings, but the required sections need to be there, with their headings. They should also be in the order listed. Please see the STAR Methods guide for more information or contact me for help.

Please ensure that original code has been archived in a general purpose repository recommended by Cell Press and that its DOI is provided in the Software and Algorithms section of the Key Resources Table. If you've chosen to use GitHub, please follow the instructions here to archive a "version of record" of your GitHub repo at Zenodo, complete with a DOI. Thank you!

Currently, you don't have a **Key Resources Table** (KRT). Note that the key resources table is required for manuscripts with an experimental component, and if a purely computational manuscript links to any external datasets (previously published or new), code-containing websites (e.g. a GitHub repo, noting that DOIs are strongly preferred), or uses non-standard software, it needs to include a key resources table that details these aspects of the paper. Purely computational or theoretical papers

that don't contain any external links and use standard software don't require a key resources table, although you're welcome to include one if you like. For details, please refer to the Table Template or feel free to ask me for help.

Thank you!

Reviewer comments:

Reviewer #2: The authors have addressed all the issues I had.
The paper is awesome and should be published in Cell Systems.

Reviewer #3: First, my sincere apologies to the authors for the long delay. My family experienced a major medical crisis with ongoing complications. My intent in agreeing to re-review had been to spare the authors adding an additional and unfamiliar reviewer, but I fear that I was forced to delay longer than anticipated. I appreciate the authors' patience and understand that they also experienced a difficult past year.

That said, I appreciate the extensive reviews and thoughtful revisions. I still have concerns that should be clarified, but do not feel the need to review this again.

The #s and headers below correspond to those from my original review.

1.1: I still don't understand 1C or what it means to say 'dashed lines are individual calcium' responses when I don't see dashed lines here and I'm not sure of the connection between 1C and 1D. Again, minor concern but why not just show the pure 0.2 Pa signal in 1C with AU and then move to R' in 1D? Regardless, I think it's essential to show when stimulation turns on/off in the plots. This was not clear originally in the plots, and I think something like the green lines in the rebuttal would help in Fig. 1C.

1.2: I understand that the authors feel that their visualization choice is "easier to grasp by the reader", but as a representative reader I have to disagree and maintain my request. I do appreciate that the authors showed a representative example of R(t) in the rebuttal, but this plot doesn't really look like the inset in 1D, so that confuses me--why doesn't it normalize to something as striking as shown in Fig. 1D's upper inset? It's fine if this is batch-to-batch variability, but I think it's important to at a minimum include these data in the supplement and explain the difference between these and 1D, upper inset.

1.3: I appreciate the rephrasing and clarification of the metrics.

* Response to request for comparison to control data and if cells actually return back to control fluorescence.
Here, the rebuttal explains how single cells relax at 70-110 seconds, while groups relax collectively at 285 seconds. The fact that cells relax to greater than their initial control value is now shown in 1D, inset and I appreciate that (and find it interesting). I do think it's important that the authors discuss where their estimates of relaxation time come from (how did they determine steady state, etc.) and include the relaxation data as a supplemental figure as it's relevant to the other figures and interpretations.

2. Information transfer and the classification types
2.1/2.2 -- I appreciate the visualization discussion and find the Fig. 5B update along with 5C to be a helpful visualization.

3. Periodic flow and memory (Figs. 3/4)
I appreciate the authors undertaking the increased rest-time experiments and clarify memory. However, there are a few things I still do not understand here.

First, I think the gap junction data are quite striking and very compelling. This is a strong datapoint supporting some form of coupling.

* Most importantly, the authors re-iterate in the text that they are "treating each round as an independent cycle, and comparing single cell responses across cycles..." My original concern was that it seems like cycles can only be independent if cells are allowed to reach a steady state between cycles, otherwise it is unavoidable that the cell will still be experiencing effects from the Nth cycle when the N+1th cycle begins. The authors appear to agree that calcium dynamics never reach steady state between cycles in the rebuttal, but then add that "cells are not relaxing rather continuously synchronizing their calcium dynamics to changing mechanical stimuli." I feel like this remains a major point of rhetorical confusion. If the cycles must be understood as truly independent, as is written in the main text (see above), then this interpretation that the cells are continuously synchronizing to a changing stimulus seems problematic for the assumption of independent cycles. I think confusion comes from the use of the words 'memory' and 'independence'. If you provide a continuously varying signal to a system and the system becomes 'better' at anticipating that signal and the standard deviation tightens, is that a sign of 'memory' or a sign of 'entrainment' (it's an excitable system that never relaxes between cycles). I do understand the argument that the cyclic stimulation data reflects a system that is better able to follow a continuous stimulus. I am probably missing something, but as not every reader will be as well versed in the correlation analyses used here better discussion in the main text is essential to clarify this.

* Regarding the experimental data shown with relaxation between cycles, I appreciate that the authors attempted this experiment and were willing to show the raw data but I am confused about this experiment was conducted. I had asked for the experiment to be run using a single pulse followed by sufficient time for relaxation followed by a new pulse, etc. Here, 4 cycles were run serially followed by a 2 minute gap (I think). In no cycle was there any response resembles that shown in Fig. 3. The lack of reproducibility is surprising—is this typical or expected and would it interfere with someone following up on this work? This also seems to have been the case in some of the other R' plots shown in the rebuttal document where they do not replicate the strength of the phenomena shown in the actual figures. Again, I appreciate the authors' transparency, but feel like this should be discussed if the system is particularly susceptible to something like passage #, time post-thaw, or individual thaws, for example.

Minor:

* Fig. 3 repeats several times that there were 13 cycles performed, but I see 15 peaks in the plots. This is relevant because I'm having a hard time understanding where the cycles stop and start and was trying to calculate cycle time. This point needs clarification. Again, including fiducial markers in all plots for when stimulation stops and starts would be a service to the reader.

* I can't interpret Fig. S8B. I'm not really sure how to read the tables, and I don't see why a correlation score of 0.3 is considered a good sign of memory. It seems like a weak behavior at best here, but again this is not my specialty. I also don't understand Table S2 at all, both due to the visual presentation (across many pages) and the vast amount of data grouped together. Again, I am likely missing something but I don't see a smoking gun for memory here.

* I think it would definitely help to provide additional context for all assumed-commonplace metrics like correlation scores. For some of us, a correlation of 0.3 is problematic, but I gather that is a good thing here, for example. There is a lot of discipline specific jargon and assumptions here and stepping back in the writing every now and then to add a sentence or two in a few key places might help improve accessibility.

Authors' response to the reviewers' second round comments

Attached.

אוניברסיטת בן-גוריון בנגב
Ben-Gurion University of the Negev
Faculty of Engineering Sciences
Software and Information Systems Engineering

Editorial Comments:

As a matter of principle, I usually only invite a revision when I'm reasonably certain that the authors' work will align with the reviewers' concerns and produce a publishable manuscript. In the case of this manuscript, the reviewers and I have a few make-or-break concerns.

1. **The reviewers raise possible alternative interpretations, which we see as crucial to address. These include:**

    A. **Reviewer 2's major point 1, raising the possibility that synchronisation may not be due to cell communication;**

In the revised manuscript, we show that gap-junctions (Fig. 3C-D) and sufficient cell density (Fig. S5A-B) are required for multicellular synchronization. These results significantly strengthen our previous results that provided less direct evidence that cell-cell communication is measureable and required for communication (Fig. 6, Fig. S9). These data exclude the concern that synchronization is a single cell phenomenon and establish that gap-junction mediated cell-cell communication is required for multicellular synchronization.

    B. **Reviewer 3's major point about how synchronisation is inferred (the section with the header 'Asynchronization');**

The reviewer was correct and our definition of synchronization in the "step" experiment is misleading and actually measures the steady state. We discarded this measurement, focused on the adaptation rate (previously termed "synchronization rate", Fig. 1E), devised a new parameterized version (Fig. S1) and devised new measurements for the heterogeneity in single cell communication (Fig. 2D). These modifications did not revoke any of the conclusions from the initial submission.

    C. **Reviewer 3's major point 3 about whether the cells had enough time to sufficiently recover between cycled to make the assumption of independence between cycles valid and to reasonably make your conclusions with regards to memory.**

We partially answer this answer in the original submission by demonstrating that the memory timescale is longer than a single cycle (Fig. S8B). We showed that there was no obvious decay in correlations up to four cycles apart (480 seconds), well beyond the timescale required for the multicellular system to relax. We performed an experiment to estimate the cells' relaxation time and another experiment with rest time between the applications of shear-stress periodic cycles. These experiments provided further evidence that the memory is independent across cycles (see response to reviewer #3 below). We decided not to include these results in our revised manuscript (reasons in the response to reviewer #3), but are willing to replicate it and include it if the reviewer and editor believe this is necessary.

2. **The reviewers request clarification of and justification for the choice of metrics and assumptions as well as choices of normalisation and representation of the data. These concerns and questions are critical to address and clarify in the next revision to ensure that the reviewers are satisfied that there are no methodological or conceptual flaws in the paper and that the claims are absolutely clear to the reviewers.**

   Following the reviewers' feedback, we made major changes in the measurements used. These new measurements are more standard and less parameterized. Here is a summary of these changes and changes that relate to other choices of normalization and representation made:

   - We excluded the measurements asynchronization and relative variability.
   - We changed the term "synchronization rate" to "adaptation rate" (Fig. 1E) and included a parameterized version (coefficient of fit to an exponential decay function, Fig. S1).
   - We replace measurements for heterogeneity (hub/individual ratio, bulk homogeneity) in communication with standardized network-science measurements, namely, local heterogeneity (Estrada index, https://doi.org/10.1103/PhysRevE.82.066102) and collective heterogeneity (https://doi.org/10.1098/rsos.160757), both depicted in Fig. 2D.
   - The response letter and manuscript (new Figures S1-S4, and Methods) include justification for parameters, metrics and normalizations performed.
   - We excluded the results regarding the enrichment beyond expectations of the communication hubs (in the pre-revised manuscript: Fig. 2F - step, Fig. 4E – cycles), because this was, at least partially, an artifact of the correlation between the transmission and receiver score, which was not specific to hubs (as reviewer #1 suspected). This has not changed any of the other results, interpretations and conclusions.

We would like to note that we decided to include many of the new results in the letter without including them in the manuscript. This is a data/analyses-intense manuscript and we prefer to include only the most critical data in it. We opted for publishing our reviews and thus these data will remain available for the interested readers.

While it will also be important to ensure that the paper is more accessible to a broader audience, and I would encourage you to have a go, I can work on this with you at a later stage.

In addition to the concerns I've detailed above, I've highlighted portions of the reviews that strike me as particularly critical. I'd also like to be explicitly clear about an almost philosophical stance that we take at Cell Systems…

We believe that data are of primary importance; text, story, and conclusions are secondary. When forced to make a choice, we would rather have text that hews closely to observations and reflects data precisely than a gloss that obscures alternate interpretations. Please keep this in mind when addressing the points the reviewers raise about choices of normalization and data presentation.

As you address these concerns, it's important that you and I stay on the same page. I'm always happy to talk, either over email or on Zoom, if you'd like feedback about whether your efforts are moving the manuscript in a productive direction. Do note that we generally consider papers through only one major round of revision, so the revised manuscript would be either accepted or rejected based on the next round of comments we receive from the reviewers. If you have any questions or concerns, please let me know. More technical information and advice about resubmission can be found below my signature. Please read it carefully, as it can save substantial time and effort later.

I look forward to seeing your revised manuscript.

Reviewers Comments:

Reviewer #1:

The authors analyze mechanosensing via calcium dynamics in endothelial cell monolayers, and characterize how heterogeneity with respect to cellular function and intercellular communication affect multicellular synchronization. The partitioning of the cell population into distinct classes of behavior (individual, leader, follower, etc.) provides an interesting way of understanding how single cell dynamics play into larger population level function.

The manuscript is written clearly, though there is a significant overlap between the methods and results sections. The figures are easy to understand, and communicate the results described in the text.

We thank the reviewer for the positive comments and enthusiasm of our results. We also thank the reviewer for raising the concern about the overlapping between methods and results. In our revision we keep in mind to strike a balance between readability and clarity. For some of the contents, such as the multiple metrics we introduced, we feel the reader would benefit from a clear definition in the main text. For others such as experimental details, we keep it in the method section. That being said, we are open to suggestions to further balance the contents.

אוניברסיטת בן-גוריון בנגב
Ben-Gurion University of the Negev
Faculty of Engineering Sciences
Software and Information Systems Engineering

MAJOR REVISIONS/MAJOR COMMENTS

The introductory discussion on heterogeneity and noise needs *significant* revision and clarity. Phenotypic heterogeneity, for example, is not the same as extrinsic noise. Extrinsic noise is the variation in expression/abundance of a species across cells in a population. It may result in phenotypic heterogeneity, but the two are not the same. It would strengthen the authors' work to distinguish these points and be more precise/deliberate throughout the manuscript, especially since the work is grounded and motivated on these important concepts.

We extensively re-wrote this part in the introduction.

"A major challenge toward establishing multicellular synchronization lays at how single cells translate environment information to intracellular signaling responses {Perkins, 2009 #1019}. Signaling in cells often rely on low copy numbers of proteins {Huang, 2007 #1460} and diffusion limited intracellular transport {Brangwynne, 2009 #1461}. These effects, often considered as intrinsic noises, lead to variable single cell signaling dynamics even in response to identical external stimuli {Swain, 2002 #1462;Elowitz, 2002 #932}. Cell-to-cell variation, or intercellular heterogeneity, is present even for cells originating from the same genetic background, also complicates our understanding of multicellular synchronization processes. Such cell-to-cell variation, or non-genetic intercallular heterogeneity, may arise from differences in gene expression levels (also terms as extrinsic noise), alternative splicing, as well as post translation modifications {Elowitz, 2002 #932;Raj, 2008 #936;Bintu, 2016 #934;Ng, 2018 #933;Gut, 2018 #1285}. Intercellular heterogeneity implies that individual cells take different states, or phenotypes, which may be induced by intrinsic noises such as by varying the copy number of receptors to modulate the probability of activation {Young, 2008 #1465}. Intercellular heterogeneity also modulates the propensity of cells to interact with their peers, as the communication between cells depend on specialized molecular channels such as gap junctions {Nicholson, 1997 #1463;Calderón, 2016 #1464}."

Following our new results on the role of gap-junctions in multicellular synchronization we also references several relevant studies:

"Others have also highlighted the role of local gap-junction mediated cell-cell communication in functional multicellular connectivity of neural progenitor cells at the vascular interface {Lacar, 2011 #1469}, in neural stem cell reactivation in the blood-brain barrier {Spéder, 2014 #1470}, in neural progenitors cell proliferation during embryonic development {Malmersjö, 2013 #1471} and in coordinated fate decisions {Ho, 2021 #1466}. To elucidate how information transfer between single cells is integrated to synchronize population-level cellular responses, we study the physiological process where monolayers of endothelial cells collectively sense and respond to external shear stress. Biologically, endothelial cells line the interior surface of blood vessels and form a monolayer that experiences varying levels of shear stress from blood flow {Yin, 2007 #1051;Hill, 2010 #1053}."

Further justification on the choice of synchronization metrics is needed. Since the current metrics are some measure of un-normalized standard deviation, the "asynchronization" measure and max-min metrics are without bounds. Classical order parameter metrics that are bound between 0 and 1 are applicable to such time-series data and would be easier to interpret.

We eliminated the asynchronization and the relative variability measurement in the "step" experiments. We kept the non-parametric "adaptation rate" (previously "synchronization rate") measurement recording the speed that the multicellular system adapts to the external stress (bounded between 0 and 1). We included a new complementary parametric model for the adaptation rate that is based on fitting of an exponent parameter (Fig. S1). We maintained the std of the cells' calcium dynamics as our measurement for synchronization in the "cycles" experiment. This measurement is sufficient because it is a relative measurement that was compared across cycles (see Table S2 for replications), also note that the raw calcium signal of each cell was normalized with respect to the cells' signal at the onset of the experiment (before shear stress was applied).

Also, since the experiments seem to be run up to 1200 seconds, the choice to use only the final 200 seconds to calculate these metrics seems odd and is not justified. In most systems, synchronization dynamics are often most "active" and relevant immediately upon application of a perturbation (rather than at the end of the experiment when the system is approaching steady state).

We agree with the reviewer and exclude this measurement as detailed above.

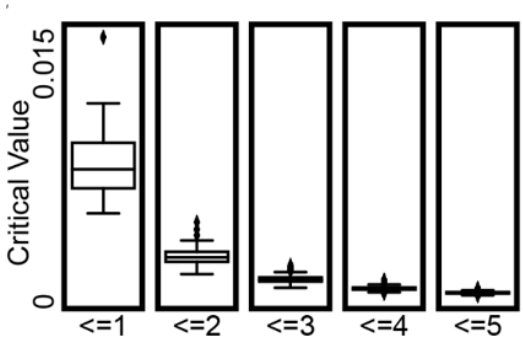
Additional major gaps (outlined below) need to be addressed/justified.

- Why are next to nearest neighbor cells included in calculating transmission and receiver scores?

The reason for using topological distances of up to two in calculating the transmission and receiver score is twofold. First, for analytic reasons and second due to irregularity in cell shape.

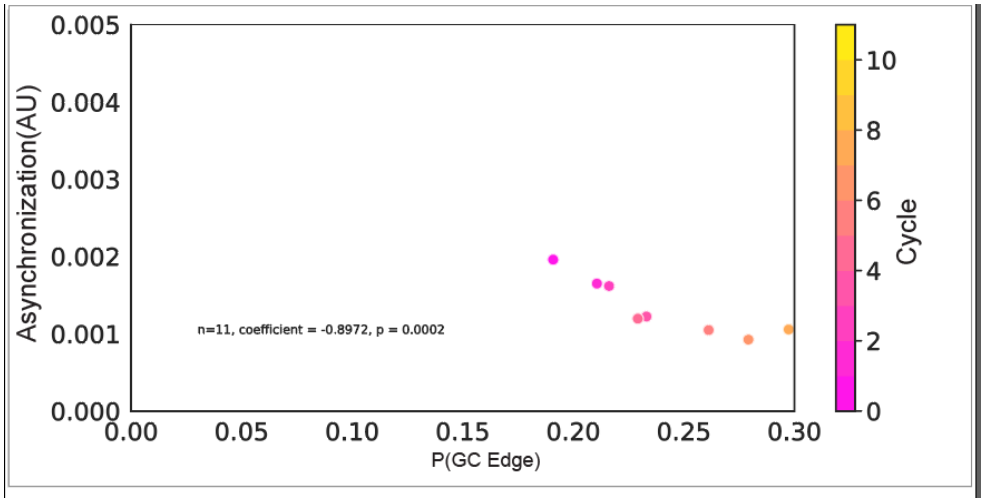
Calculating the transmission and receiver score based on a few neighbors will increase the uncertainty. Topological distance of two is a sweet spot in terms of reducing the false-positive errors. We specify this reason in the methods, "These neighborhood sizes were determined for sufficient observations for statistics, and the expected short-range communication between the cells." And now produced supporting data to make this point quantitatively.

Specifically, increasing the number of neighbors makes the statistical test for determining GC edges more strict (from ~0.008 for topological distance of one to ~0.003 for topological distance of two) because we use Bonferroni correction to determine statistical significance. This way, the statistical test also becomes less variable between cells since the variability in the number of neighbors also decreases. This new analysis is shown in the new Fig. S2B (see below).

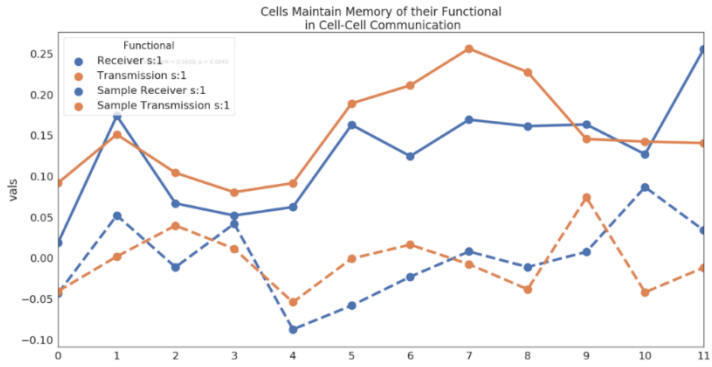


Importantly, all the main results regarding heterogeneity, memory, and local-to-global information propagation hold when considering topological distance of one (see below, not included in the revised manuscript).

Gradual increase in information flow and synchronization over time:

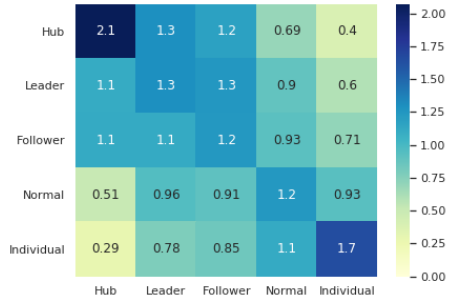


Memory:

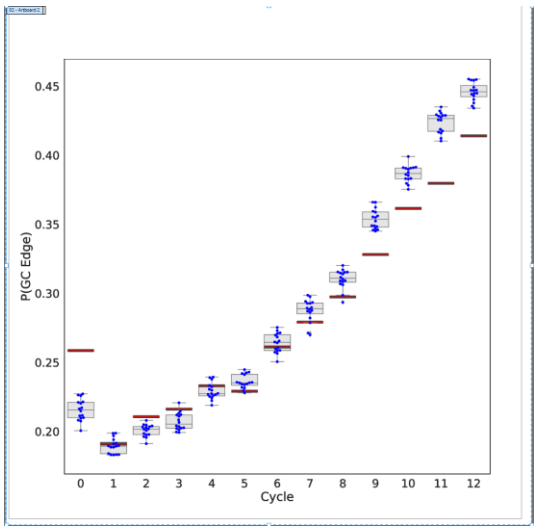


אוניברסיטת בן-גוריון בנגב
Ben-Gurion University of the Negev
Faculty of Engineering Sciences
Software and Information Systems Engineering

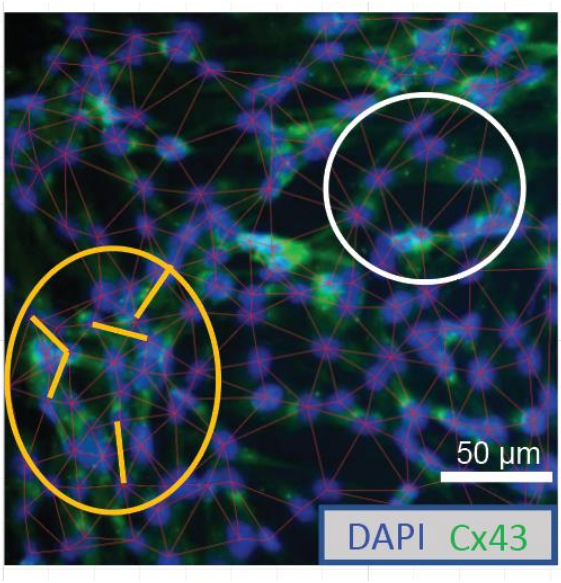
Enrichment factor of cellular state transitions



The observed versus permuted Granger causality edge probability, P(GC edge), over the cycles



A second reason for choosing to include next-to-nearest neighbor in the analysis due to the irregular shapes of the cells. To illustrate this, we analyzed a typical immunofluorescent image (see below) where cell nucleus (blue), and gap junction proteins (green) were labeled simultaneously. The red lines connect nucleus of nearest neighbor cells that were defined with Delaunay Triangulation. In the low density region (such as the area within the white circle), gap junctions only appear between nearest neighbors. In the high density region (such as the area within the yellow circle), gap junctions also connect cells that are next-to-nearest neighbors (see example pairs indicated by yellow lines). Our shear-stress experiments were mostly performed at high cell density, with confluence greater than 85%. Therefore it was appropriate to include cell pairs with topological distances equal to two. This new analysis is shown in the new Fig. S2C:



- What might be a biological/physical motivation to go beyond a topological distance of 1?

Because the cells take very diverse and irregular morphology, nearest neighbors defined by Delaunay Triangulation may be too restrictive. Therefore, we include next-to-nearest neighbors to make the counting more robust. See above.

- It seems that counting next to nearest neighbors while calculating these scores presents double counting with respect to intercellular communication. How do the authors mitigate this bias?

Multiple cells can communicate with one another concurrently. We only consider unique pairs in calculating the GC edge probability for each cell, therefore there is no double counting in term of considering a potential edge more than once. Our method relies on local pairwise analysis of cells, and defines single cell communication properties without requiring explicit construction of the network or committing to a specific network architecture. This latter point is mentioned in the Discussion.

It is true that a limitation of calculating pair-wise Granger Causality is the identification of indirect causal effects (e.g., cell A influences B that influences C, and an A→ C GC-edge). This limitation is also relevant when considering only nearest neighbors (i.e., neighborhood defined by topological distance = 1). Our analysis showing marginal correlation between the cells' neighborhood size (at topological distance ≤ 2) and the corresponding degree rank (Fig. S2D), no correlation between the neighborhood size and the transmission/receiver score (Fig. S5 in the initial submission, excluded from the revised submission, see below), along with the clear biological signal across cycles and after perturbations imply that this concern do not affect any of our conclusions.

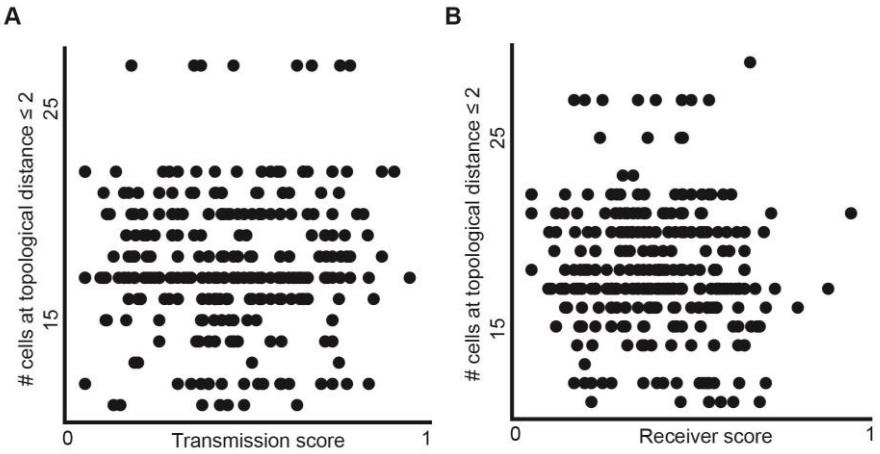


Figure S5: The number of neighbors at topological distance ≤ 2 is not correlated with the transmission or receiver scores. N = 295 cells from one experiment. (A) Transmission score. Pearson coefficient = 0.04, p-value = 0.47. (B) Receiver score. Pearson coefficient = -0.09, p-value = 0.11.

How might intercellular communication be realized in this context?

Our analysis is based on defining measurements for single cell communication properties. Our additional experiments show that multicellular dynamics, particularly the synchronization, are significantly impacted by gap junction inhibitions, or by reducing cell-cell contact with sparser density. Therefore, the evidence support gap junctional communication.

- While the GC scores provide a statistical basis to suggest it exists, how might such communication happen physically/chemically? Connecting it back to mechanism--even if hypothetical--would help ground the study.

Based on previous reports from ours and other groups, we suspect there may be two channels for intercellular communications. One is the exchange of messengers such as IP3 through gap junctions, the other is the through stress-induced release of ATP molecules that diffuse in the extracellular media. The latter is unlikely because we do not observe GC links preferably form in the direction of flow. Our experiments with gap junction inhibition and sparse cell seeding (Fig. 3C-D, Fig. S5A-B – see these results in response to reviewer #2) provide strong support for a communication mechanisms that is driven by direct cell-cell contact via gap-junctions. This is now a major result in the revised manuscript.

The use of z-scores should be justified since the distribution of GC coefficients seems to be asymmetric, both for transmission and receiver scores.

We use z-score normalization to account for the experiment-to-experiment variations such as cell density and any other batch effects. This is now justified in the Methods ("The transmission and receiver scores of each cell were normalized across the population to allow direct comparison of single cell heterogeneity between cycles and between experiments.", "We calculated the receiver

and transmission z-score for each cell c_i, the variation from the mean in units of standard deviations: Tr_norm(ci) = (Tr(ci)-μ)/σ, where μ is the mean transmission score across the population over all cycles, and σ is the standard deviation.") and in the main text ("The normalized scores allowed direct comparison of the single cell heterogeneity across different cycles in the same experiment as well as between different experiments."). Z-score normalization was used to empirically assign cells to different communication states. The normalization was performed for the receiver and for the transmission score independently. An alternative normalization scheme would yield similar qualitative conclusions.

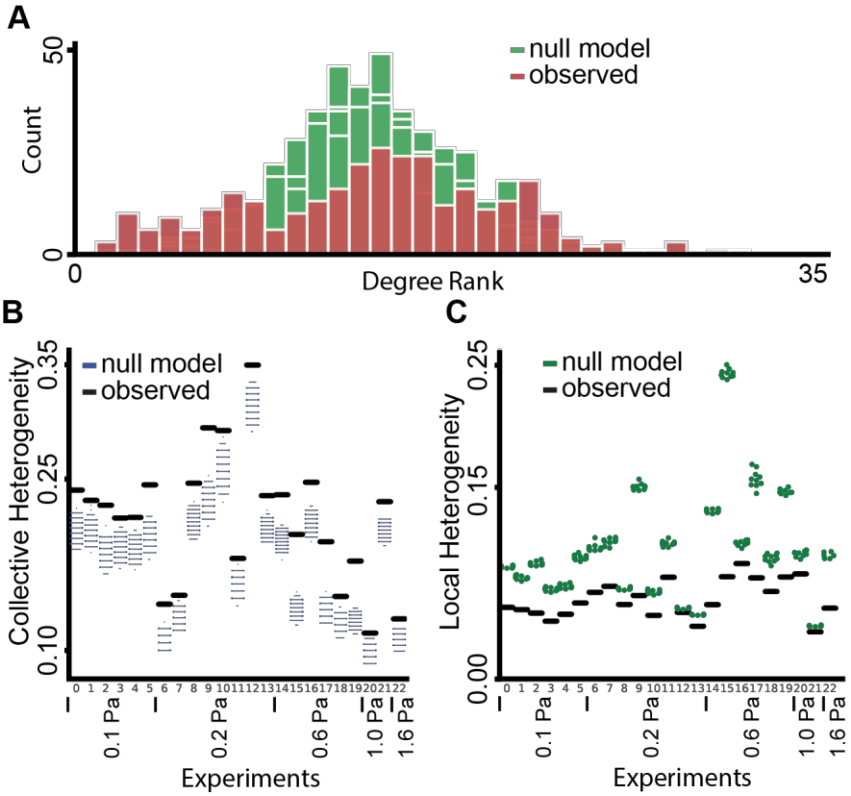
What could one infer from the fact that the observed fraction of hubs is greater than P(z_tranmission >0.5)*P(z_receiver >0.5) other than the two scores being non-independent? Why might a cell capable of receiving information also be more likely to transmit information and vice versa?

Following the reviewer's comment we assessed the relation between the transmission and receiver score and found that they were correlated (see table below). The fact that transmission and receiver scores are related, provides further indication for non-random self-organization in the multicellular network. However, we can not exclude the possibility that the "enrichment of hubs beyond expectation" (Fig. 2F and Fig. 4E in the initial submission) is a mere outcome of this correlation so we decided to exclude it from the revised manuscript. This has not changed any of the other results, interpretations and conclusions.

	cycle	corr	p_value
0	0.0	0.350112	1.700789e-20
1	1.0	0.302209	1.997584e-15
2	2.0	0.239153	4.737508e-10
3	3.0	0.191850	6.717544e-07
4	4.0	0.182960	2.190592e-06
5	5.0	0.302384	1.921515e-15
6	6.0	0.210406	4.747027e-08
7	7.0	0.275393	5.730064e-13
8	8.0	0.211083	4.289757e-08
9	9.0	0.192519	6.132150e-07
10	10.0	0.127612	1.008504e-03
11	11.0	0.170090	1.098633e-05
12	12.0	0.099001	1.087309e-02

In calculating the EMD metric for bulk heterogeneity, why is a uniform distribution used as the baseline? What might suggest an equal distribution of roles in the population/why might it be a useful prior?

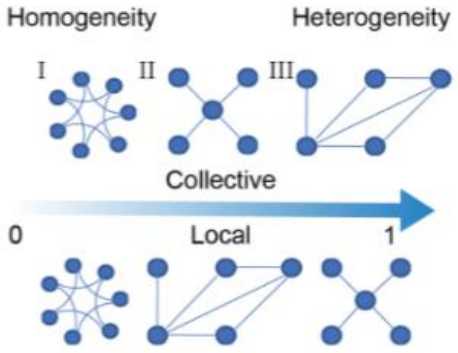
We agree with the reviewer that the uniform distribution, although simple and intuitive, might not the optimal baseline. We now use simulated distributions either simulating randomly the edges for each experiment based on the edge probability of that experiment as a null model for collective heterogeneity or by shuffling the edges while preserving the degree ranks as references for measuring local heterogeneity (new Fig. S3 – see below).



In Fig. 2E, the authors mention that "These results suggest that increased fraction of communication hub cells along with decreased fraction of individuals were associated with an improved synchronization process". It seems reasonable that when more cells participate in communication and fewer take on an exclusively individual role, a more coherent outcome is likely. Isn't this expected?

It is indeed expected, and aligns with our aim to highlight the role of (local) cell-cell communication in (global) synchronization. Following the reviewers comments we excluded the hub/individual ratio from the manuscript. Instead we make the same point by using more standard measurements for heterogeneity (new Fig. 2D, see depiction below), and with our result

regarding the increased fraction of communication hubs over cycles (Fig. 5B) and the stability of the communication hub state (Fig. 5C).



MINOR REVISIONS, SUGGESTIONS & COMMENTS

The classification of individual/common/leader/follower/hub cells happens when the scores are either <-0.5/(-0.5,0.5)/>-0.5. A negative sign needs to be added where appropriate.

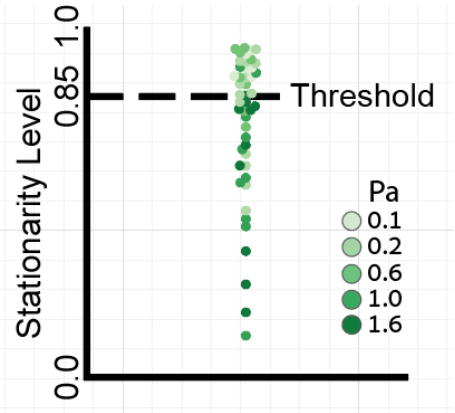
Done

Replace "stableness property" with "stability".

Done

A summary of how many cells (as a fraction of the whole) passed both stationarity tests in each experiment would better support the results.

This information is now available in Table S1 and shown in Fig. S2A. See below:



Since heterogeneity can take on many forms, it would be useful to anchor the use of the term to a specific context throughout the manuscript and to use them in a better defined and deliberate manner.

In the revision we defined two terms, local heterogeneity (Estrada index) and collective heterogeneity. These are standard measurements to assess degree rank heterogeneity and are now used throughout the manuscript.

What could be an explanation for the sharp increase in the hub to individual ratio in the last two cycles?

This could be due to positive feedback in the hubs (increased slope) that leads to faster growth, together with stable linear decrease in the individuals, which count for more as the number decreases. For the revision, we switched to other measurements for heterogeneity, as discussed above, so the hub/individual ratio is not considered any longer.

Reviewer #2:

The paper by Zamir et al addresses an interesting question related to the synchronization of Ca2+ signaling among a population of endothelial cells. The questions raised are interesting and the overall approach used in the paper is appropriate. The results claim to show how synchronization initiates as local communication that propagates toward the population as a whole. If correct, these results are of interest to the broader community, and the paper a great fit for Cell Systems. However, there were a few things that confused me about this paper's methodology and analysis. These confusing points have to be addressed before I can endorse publication. If the author could better support their claims and clarify the key points raised below this paper will be a great candidate for publication at Cell Systems. However, there is a chance that the confusion is not just a misunderstanding but points to a logical flaw and if that is the case the paper should not be published. My recommendation is to return this to the authors and allow them to respond. Note that no additional experiments are requested.

Major points.

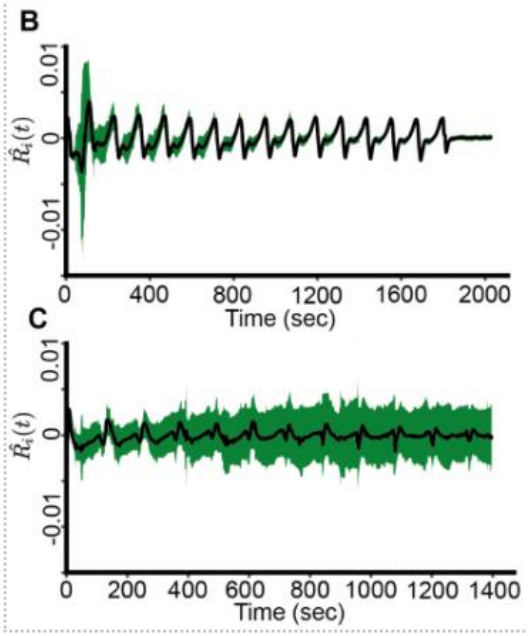
1. Synchronization might not be due to communication.

This is a tricky one and my main concern about the interpretation of the results. Let's say that cells are not communicating at all. But, each one of them, independent of the others, is all getting better at expecting a spike, i.e. change in flow rate. This could be something simple like a negative feedback loop on receptor occupancy, intracellular regulation of some kinase that regulates ATP receptors, or something like that. In that case, all cells will still look like they are synchronizing even though they are not communicating with each other. All the global measures could still look similar to figure 1 and 3BC. In that case, since cells are more coordinated, they should also be a bit more predictive of each other and I can imaging that the GC edges might still be significant even if they are not really

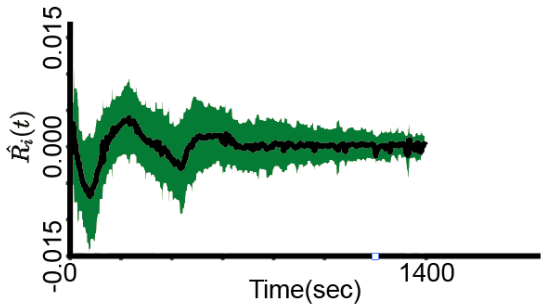
talking to each other. They are significant because other cells (that are becoming more correlated with each other) will carry some information on the cell's future response. What I was missing is an analysis that shows that all (most) the GC-edges disappear when you do spatial permutation. Yes, I am aware of figure S8A but it shows the data in a compounded way (since it takes into account the hub/individual definition) so it's harder to interpret. In my understanding, any GC-edge that is still there after spatial permutation is not "real" communication. If there are a lot of these after such permutation, how can one interpret these edges as truly showing information flow? This is clearly shown in figure 5A where the bulk of the changes in P(GC-edge) is easily explained by the permuted data and only a tiny fraction is really left (see inset). If indeed cell cell communication is the key factor in creating synchrony, then the difference from random (spatially permuted) should be much much bigger.

We addressed this critical concern both experimentally and via further analysis.

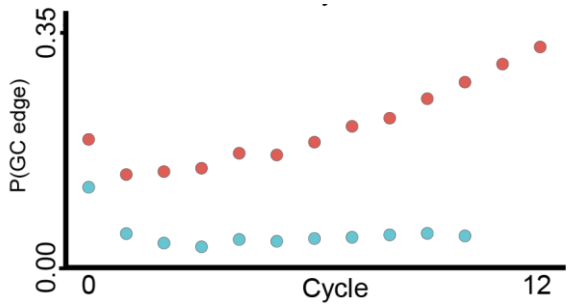
The multicellular system did not synchronize in response to gap-junctions inhibition (Fig. 3B versus Fig. 3C, see below, all results are available in Table S2).

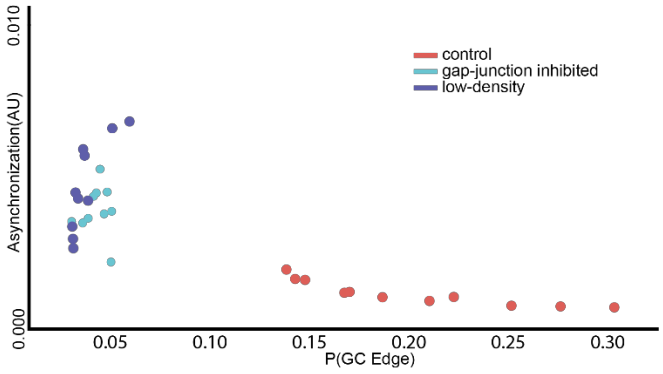


With sparser cell seeding the multicellular system did not synchronize as well (Fig. S5A, see below, all results are available in Table S2).



In both perturbations, information flow was lower and did not increase in time (new Fig. 3E – see below (red - control, cyan - gap-junction inhibited), new Fig. S5, Fig. S6 – see below, Table S2).





These results show that gap-junctions and physical cell-cell contacts are required for this mode of cell-cell communication and with multicellular synchronization. These results align with previous work from Sun's lab on Fibroblast cells (PMID: 22566661, PMID: 27573834).

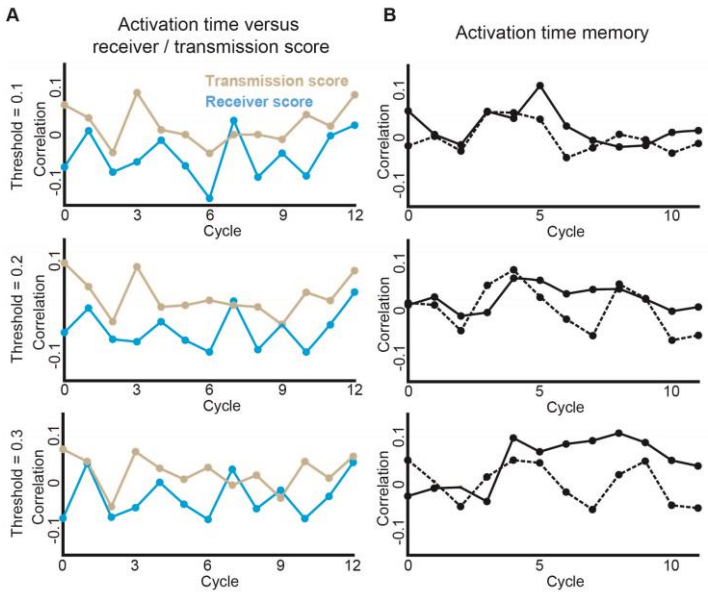
We refer to this point explicitly in the Discussion:

"We showed the cells were actively communicating with one another locally, and that physical cell-cell contacts via gap-junctions were required for multicellular synchronization. These

conclusions were supported by multiple lines of evidence throughout our study. First, we reported that gap-junctions and sufficient cell confluence were required for multicellular synchronization (Fig. 3B versus Fig. 3C or Fig. S5 and Table S2). This results provided the strongest evidence against the concern that the multicellular synchronization was caused by differential cell autonomous response to the external shear stress that was applied on the group. Second, we demonstrated that both local and collective heterogeneity depended on the spatial organization of cells in their vicinity (Fig. S3). Third, we found that the activation time, a cell's autonomous response to the external stress, was not associated with the transmission or receiver score (Fig. S9A), which would also be conflicting with the enrichment of communication hubs (which are both leaders and followers). Forth, cells "remembered" and reinforced their roles in the multicellular communication network over time, as a local, spatially-dependent property (Fig. 4C), but did not "remember" their activation time in previous cycles (Fig. S9B). Fourth, neighbor pair cross correlation was a local cell property throughout the experiment (Fig. 6A). Together, our data established the decoupling of the local cell-cell communication from the global external stimuli, and established that the emergence of multicellular synchronization required gap-junction mediated local cell-cell communication."

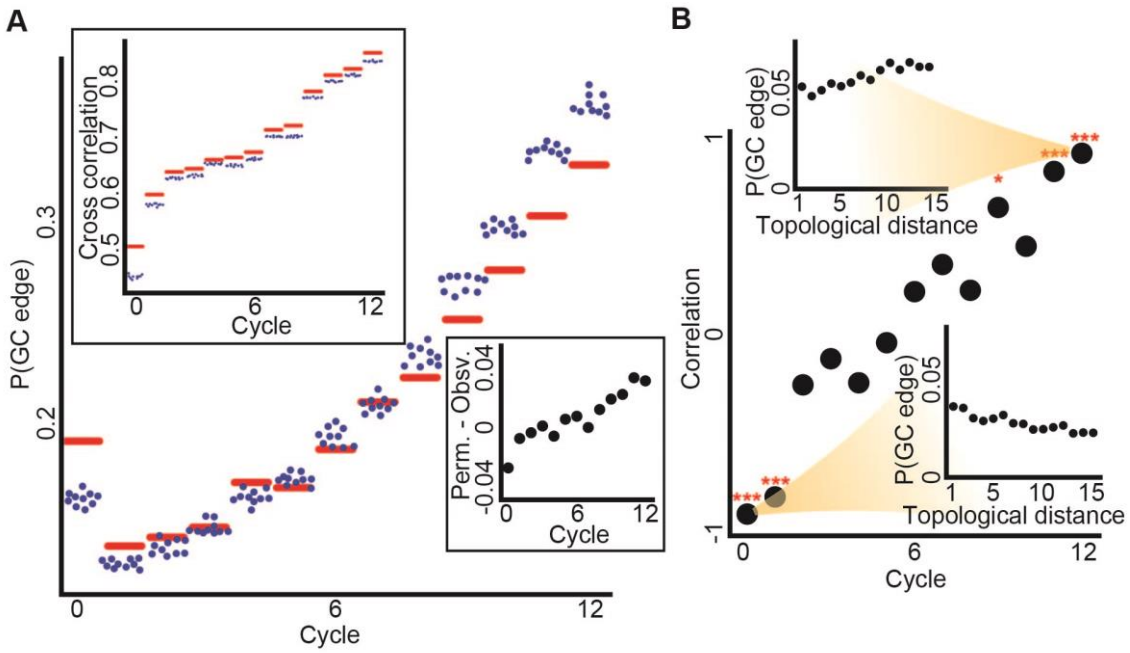
These results provided the strongest evidence against the concern that the multicellular synchronization was caused by differential cell autonomous response to the external shear stress. Thus resolving the reviewer's concern that "**each one of them, independent of the others, is all getting better at expecting a spike, i.e. change in flow rate**".

Moreover, this concern is also partially rebutted with our analysis showing that communication properties (receiver/transmission scores) were not correlated to the autonomous cells' response to the external mechanical stimuli as measured by their activation time (Fig. S9 - see below).



אוניברסיטת בן-גוריון בנגב
Ben-Gurion University of the Negev
Faculty of Engineering Sciences
Software and Information Systems Engineering

The reviewer argument "**since cells are more coordinated, they should also be a bit more predictive of each other**" is challenged by the earlier presented data (perturbation experiments and response time to the external signal) with more support provided by the data we present in Figure 6 (see below). We found that once neighboring cells reach sufficient synchronization their ability to influence each other (as measured with Granger Causality) is less effective than cell pairs far apart, which are less synchronized. In the upper inset in panel A we showed the cross correlation between neighbors (red), and after spatially permutations (blue). The correlation increased with time implying improved local synchronization/coordination. The spatial permutation always (slightly) decreased the correlation, indicating that cross correlation was maintained as a local cell property throughout the experiment. In contrast, in the main panel we showed that the spatial permutation decreased the edge probability in early cycles but increased it in later cycles (see the lower inset transitioning from negative to positive values over the cycles). These results suggest that once cells are synchronized beyond some level with their neighbors, they become more probable to communicate with more distant cells, which they are less synchronized with. This result is then verified in panel B.



The reviewer arguments "**missing is an analysis that shows that all (most) the GC-edges disappear when you do spatial permutation"** and "**In my understanding, any GC-edge that is still there after spatial permutation is not "real" communication.**" are quite tricky because we show (in panel B above) that GC edges are not limited to short distances, in fact, long-range edges become more probable as the system becomes more synchronized locally (and globally). This implies that the long-range communication, or the multicellular synchronization, is a confounding factor masking the true local components in the network. This confounding factor is less drastic in the first cycle, where the lower bound of the local contribution of communication is ~0.03 from the observed 0.19 (i.e., over 15%).

Ben-Gurion University of the Negev, Be'er Sheva, Israel. Email: assafza@bgu.ac.il, lab website: https://www.assafzaritsky.com/

2. Information flow - significance vs effect size.

Another thing that confused me is the use of statistical significance to decide if a GC-edge is there rather than the actual effect size. Two cells can communicate with each other in a way that barely influences their Ca2+ dynamics, yet the effect is not random, i.e. p-value will be below the threshold. One of the reasons I suspect that the p-values are so small is that you are treating each timepoint as independent in the way you set up the F-test. This means that the sample size is very large and a very tiny effect could still be called significant. Defining cell to cell communication based on effect size rather than p-values will be more appropriate and could be informative in explaining how much cells really influence each other. This might compound the interpretation of the GC-edges.

Our perturbation experiments (gap junction inhibition, sparser cell seeding) drastically reduced the P(GC edge) and synchronization (Fig. 3, Fig. S5). We believe that these results provide sufficient evidence for the importance of local cell-cell communication regardless of any confounding factors in the current measurements.

Importantly, we take a very conservative approach by using Bonferroni which sets the significance threshold to 0.05/n where n is around ~15-20 (Fig S2), it means a statistical test threshold between 0.0025 - 0.0033 which is quite strict.

As described in the Methods: “To fix spurious edges due to multiple hypothesis testing we applied the strict Bonferroni correction that defines the edge significance threshold based on the number of edges considered .In our case, with a significance threshold of 0.05 and n - number of potential edges we get a new significance threshold of 0.05/n”

As for the reviewer’s response “**One of the reasons I suspect that the p-values are so small is that you are treating each timepoint as independent in the way you set up the F-test. This means that the sample size is very large and a very tiny effect could still be called significant.**”, the F-test takes into account the sample size and penalizes longer time series with stricter statistical significance. To provide empirical evidence we calculated the (variance of the errors in the restricted model - variance of the errors in the unrestricted model)/(variance of the errors in the restricted model) as a fraction of the error reduction. This normalized measurement for the magnitude of the influence was correlated with the Granger Causality p-value (Pearson correlation coefficient for the representative cycles experiment = -0.59).

3. Cellular "roles" definition.

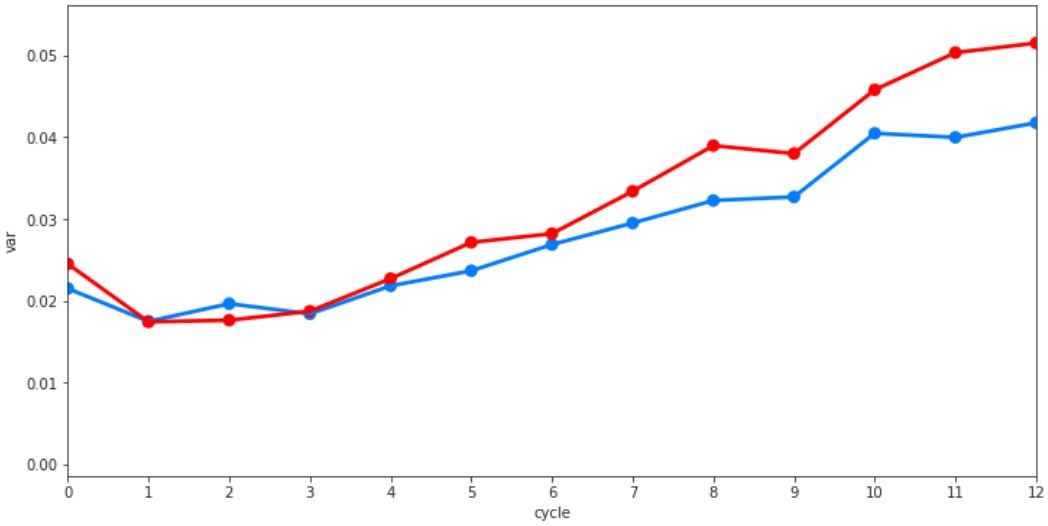
I am confused about the need to normalize the transmission/receiver score and why the normalized version is used to define the different "tasks" cells putatively take. As cell become more uniform in their response over activation cycles, the variance will shrink and the z-score will be inflated. I understand that there is a lot experiment to experiment variability and therefore the author felt some need to normalize, but I think that z-score

could mess up the interpretation as the variance is not expected to stay the same across cycles. The change in the spread of points in Figure 3E could easily be related to normalization and not underlying. I'm not sure exactly how, but a version that takes into account spatial permutations for normalization seems more appropriate to me.

In the "cycle" experiments, the z-scores are calculated based on a full experiment (mean and variance inferred from the integrated data over all cycles) and then compared across cycles. Thus, there is no concern about normalization artifacts, specifically, the variance between cycles is highly correlated before and after normalization.

We used z-score normalization to account for the experiment-to-experiment variations and to empirically assign cells to different communication states. Performing no normalization or alternative normalization schemes would yield similar qualitative conclusions.

To assess the reviewer's concern regarding "**As cell become more uniform in their response over activation cycles, the variance will shrink and the z-score will be inflated**", we calculated the variance in the cells' transmission (red) and receiver (blue) scores for each cycle without normalization and found that it increased over time, see below. This result also supports our results associating collective heterogeneity and synchronization. We believe that the normalization across the full experiment rather than for each cycle independently resolves the reviewer's concern.



Minor/stylistic comments - some of these points are a matter of opinion and I try to point out which is which.

1. (opinion) Reduce the number of measures defined to focus on the ones that matter most. Some of these seem pretty arbitrary and the correlation between them (for example, sync rate vs rel var) are true almost by definition.

Agreed. We reduced the number of measures and focused on a more standard set of readouts that essentially measure different properties.

2. In a few places, you compared P(hub) vs P(z-transmissoin>0.5) & P(z-reciver>0.5) isn't that how a hub is defined in the first place?

Not exactly, for a cell to be defined as a hub both the normalized transmission and the receiver score are above 0.5. This implies that if the transmission and the receiver scores are intendent than we would expect P(hub) = P(trans > 0.5) * P(rec > 0.5). Our analysis showed that communication hubs are enriched beyond what one would expect from independent marginal distributions. However, following reviewer #1's concern that this "unexpected enrichment" could be a mere outcome of this correlation we decided to exclude this result from the revised manuscript.

3. (opinion) the abstract was missing a sentence of two that frame the question of the paper. The first few sentences go straight to describe what was done and what was found. It would be better to explain why and place this is context.

We now start the Abstract as follows "Multicellular synchronization is a ubiquitous phenomenon in living systems. We quantitatively characterize how noisy and heterogeneous behaviors of individual cells are integrated across a population toward multicellular synchronization by studying the calcium dynamics in mechanically stimulated monolayers of endothelial cells.".

4. (opinion) in the discussion the first paragraph is a bit broad and is not sufficiently focused on summarizing the results. A lot of that is happening in 2nd paragraph instead. Some of the content of the 1st paragraph that describes the broader context ("This model-free data-driven approach can be applied to a broad set of biological systems…") should come later-on.

The first paragraph described the methodological approach and its potential. We prefer to keep it in tact.

5. The reason behind the choice to use derivative instead of directly using Ca2+ value was not fully explained.

For all the correlation type of analysis, ours included, differentiation is necessary for the following reasons. First, both exchange of IP3 and Ca lead to a change of calcium concentration, so the derivative is quantity directly related with communication. Second, the derivative has nice properties such as independent of basal level of signal, independent of slow systematic errors such as photo bleaching, and stationarity. We added this explanation to the text and cited other papers that used the signal derivatives ({Sun, 2012 #255;Sun, 2013 #1027}).

אוניברסיטת בן-גוריון בנגב
Ben-Gurion University of the Negev
Faculty of Engineering Sciences
Software and Information Systems Engineering

Reviewer #3:

Summary: This paper poses a really fascinating question about the extent of signaling information propagation and correlation within a responsive and heterogenous tissue. This is quite different from other studies I have seen that primarily emphasize mechanical force coupling, and it tackles the problem of heterogeneity head-on. It lays out a series of metrics and approaches to try to break down and quantify the coupling effects using approaches from statistical mechanics and information theory, and the level of quantitative analysis seems excellent and thorough. However, I confess that I found it difficult to follow at points, especially justifications for why specific assumptions were made, and for what I felt was a lack of sufficient explanation for a broad audience. I also had some specific technical concerns about the quantifications and experimental approaches that I think the authors can easily address. Overall, I think this is an exciting and important paper for the field.

My comments below track the order of the figures.

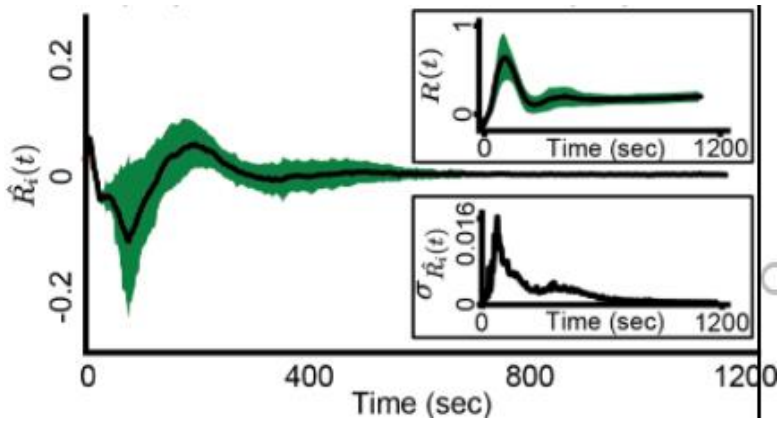
1. Figure 1 and establishment of baseline measurements

The microfluidic shear assay is well conceived at a hardware level, but the data presentation is somewhat confusing.

i. Stepped pressure vs. fixed pressure

I understand that stepped flow speeds can be generated (Fig. 1B) and that this results in changed in the calcium indicator intensity (1C), but I think the authors need to show clearer data for 1C of just 0.2 Pa in single step configuration rather than showing 0.20.3. If all of the paper is conducted at 0.2 Pa, show us only 0.2 Pa held for 300 seconds and scaled so the dynamics are clearer.

The experiment shown in Fig. 1 meet these criteria. In the revision, we also included the raw calcium signal, see below in the new upper inset of Fig. 1D:



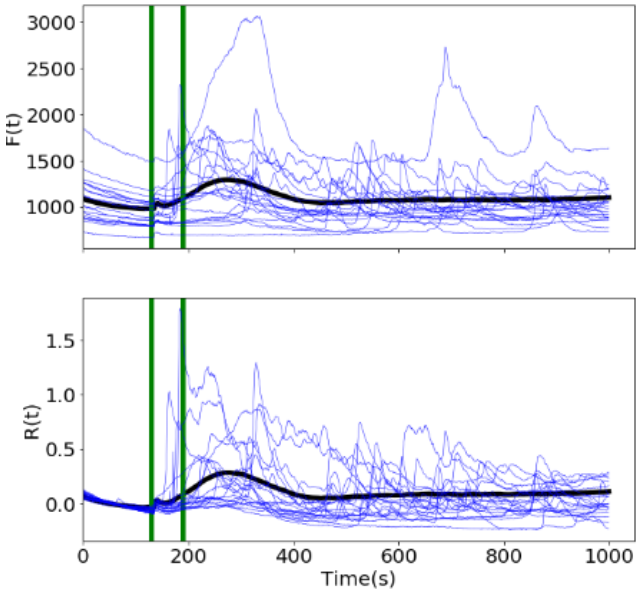
Ben-Gurion University of the Negev, Be'er Sheva, Israel. Email: assafza@bgu.ac.il, lab website: https://www.assafzaritsky.com/

ii. R'_i and standard deviation presentation (1D/E)

Where does stimulation start? Is t=0 when the 0.2Pa pressure shift is first applied? If so, I think the plots should be redone to show control data (without any shear stimulation) and then clearly indicate where the shear perturbation starts. I would show control data going back at 1 min. I know that R(t) normalizes by F_o, but I'm also confused about why R(t) starts at a non-zero value, and I think this has to do with the control period pre-stimulation not being shown.

We have this experiment shown in Fig. 1C where the 0.2Pa pressure starts ~20s seconds after the imaging onset. In the rest of the results t = 0 is the time when shear stress is first applied. We feel that this visualization is easier to grasp by the reader.

For the reviewer we provide an example starting before shear stress is applied. The top plot shows the raw intensity F(t) of 20 randomly selected cells (black – mean, green vertical line marks the time of the stimuli onset and stop (130s idle → 60s shear flow → idle). The bottom plot shows the normalized signal R(t) with values ~0 before the application of shear stress.



iii. The three metrics

The attempt to carefully define synchronization metrics is critical, but I had some confusion about why these specific metrics were chosen and how the assumptions were justified (e.g. 200 sec windows, etc.). I think the main text itself needs to have a more complete discussion of the metrics and why they are useful/what assumptions were made. I also did not find the additional discussion in the Methods to clarify the metrics.

We agree with the reviewer and revised accordingly with more standard and less parametrized measurements. These include excluding the asynchronization and relative variability measurements in the "step experiments", and replacing the hub/individual ratio and bulk homogeneity with collective (https://doi.org/10.1098/rsos.160757) and local (https://doi.org/10.1103/PhysRevE.82.066102) heterogeneity.

*** Asynchronization:**

o Why pick the last 200 seconds instead of more objectively defining steady state?

o The implication seemed to be that a low Asynchronization implied synchronization (opposite of asynch. is synch.), but it seems what it actually implies is that many cells have settled down and have minimal rates of fluorescence change, which doesn't seem to imply that those cells are necessarily coupled, just that their fluorescence intensity is no longer fluctuating (leading to low R'(t) and lower standard deviations. I am likely misunderstanding something, but it would help to have a better discussion of this in the next.

*** Relative variability**

o See above—not clear to me as a reader why the lower standard deviation inherently implies synchronization.

o Max standard deviation - min standard deviation seems to capture the total spread of the variation, and that seems important, but I don't understand what this is 'relative' to since it's not a dimensional quantity. Does this make more sense as a ratio, or normalized with respect to control data?

We agree with the reviewer regarding the arbitrariness of the 200 seconds and that the cells have relaxed and thus excluded the asynchronization and the relative variability measurements. We do use asynchronization in the "cycle" experiments because there the cells are experiencing continuous changing shear flow and so their fluorescent signal keeps fluctuating.

*** Synchronization rate**

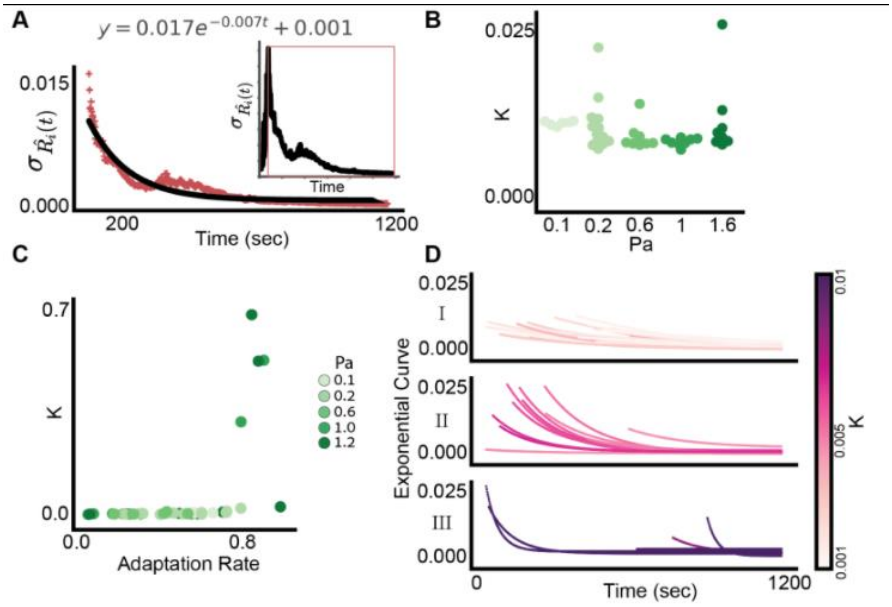
o Why 200 seconds post-peak standard deviation? Could you instead measure the time for the system to go from peak deviation to minimum deviation?

Following the reviewer previous clarification we changed the term "synchronization rate" to "adaptation rate". We chose a temporal window of 400 seconds (200 frames) because this was the maximal time-frame without excluding experiments (several experiments had late peak times defining an upper bound on the temporal window – see details in the following table). This reasoning is now explicitly described in the Methods.

אוניברסיטת בן-גוריון בנגב
Ben-Gurion University of the Negev
Faculty of Engineering Sciences
Software and Information Systems Engineering

experiment number	total time	end of window	peak time
0	1170	698	298
1	1170	434	34
2	1170	1152	752
3	1170	504	104
4	1170	652	252
5	1170	554	154
6	1170	566	166
7	1170	656	256
8	1170	528	128
9	1170	484	84
10	1170	818	418
11	1170	478	78
12	1170	824	424
13	1170	922	522
14	1170	878	478
15	1170	594	194
16	1170	436	36
17	1170	904	504
18	1170	790	390
19	1170	858	458
20	1170	654	254
21	1170	780	380
22	1170	436	2

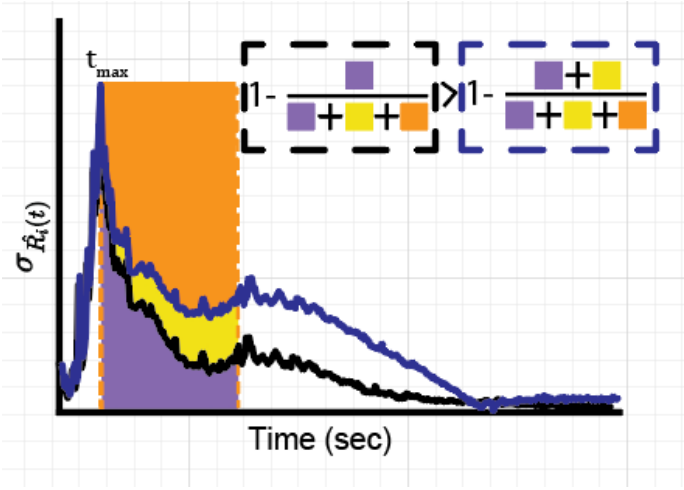
To resolve concerns regarding the post peak time parameter, we included a parameterized version for the adaptation rate defined as the coefficient of fit to an exponential decay function (Fig. S1, see below). One caveat is that some experiments did not fit well to this project. Still, the correlation to the non-parametrized measure was high (Fig. S1C).



o I did not understand the definition of the ratio of the integral over 200 s post-peak to the 'theoretical upper bound'. The figure and text were difficult to follow in this regard.

Ben-Gurion University of the Negev, Be'er Sheva, Israel. Email: assafza@bgu.ac.il, lab website: https://www.assafzaritsky.com/

We revised the text and included a more clear illustration (see revised Fig. 1E).



Revised text: "We defined the *adaptation rate* as $1-\left(\frac{\int_{t_{max}}^{t_{max}+400}\sigma_{\widehat{R}_i(t)}(t)}{400*\sigma_{\widehat{R}_i(t_{max})}}\right)$, where $\sigma_{\widehat{R}_i(t)}(t)$ is the population-level standard deviation of single cell calcium dynamics $\widehat{R}_i(t)$ at time *t*, and $t_{max}=\underset{t}{argmax}\,\sigma_{\widehat{R}_i(t)}(t)$ is the time of the peak variability in calcium dynamics (Methods, Fig. 1E). The adaptation rate is a non-parametric measurement for the speed that the multicellular system adapts to the external stress. When a system rapidly synchronizes, adaptation rate approaches one. Conversely, if a system maintains a large deviation between the dynamics of individual cells, adaptation rate is close to zero. In general, higher adaptation rate implies faster multicellular adaptation to the external stress (Fig. 1E black curve adapts faster than blue curve)."

o What units does this have that make it a rate?

The non-parametric adaptation rate measure is normalized between 0 (no adaptation) and 1 (rapid adaptation).

*** Additional parameters**

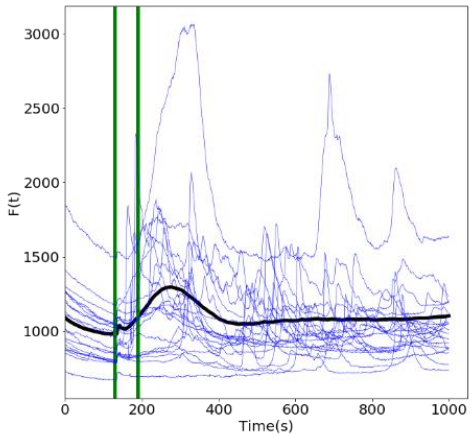
It seems like it would help to compare or normalize to control data. It would also help to understand if the cells ever actually truly equilibrate back to their control fluorescence intensities.

The adaptation rate is a relative measure with normalized values between 0-1 measuring the speed that the multicellular system adapts to the external stress.

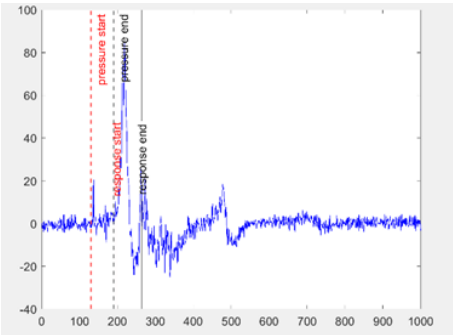
To assess single cell relaxation time we performed a new experiment (2 replicates) where we applied a pulse of shear stress (0.2 Pa) for 60 seconds and then released the stress and followed

the relaxation. We found that single cells reached a steady state ~70-110 seconds after their peak (mean 72.25 seconds, standard deviation 25.3 seconds) and full collective relaxation was achieved after approximately 285 seconds, as detailed below.

Shown below are 20 randomly selected single cell raw calcium intensity time-profile. Black: mean intensity. Green: Shear stress application starting and stopping time.



More specifically, single cell analysis was performed by selecting 20 responding cells, manually marking their calcium peak near the time to the pulse of shear stress application and the time where the cells settle. See below a representative cell, the red solid line indicates the beginning of the pulse and the black solid line indicates the end of the pulse. The fluorescent intensity usually do not return to the pre-stimuli level, but the cells settle to a new basal level that is slightly higher than the original, and do not show changes in the intensity beyond short noise fluctuations (see Fig. 1C, and the new upper inset in Fig. 1D). This is one of the reasons we consider the derivatives in our analysis. We do not feel that this data is necessary for the manuscript and prefer not to include it in the already data-dense manuscript (unless the reviewer finds it critical).



o I think it would help to show the ratio of sigma_steady_state to sigma_ctrl to clarify this. If the system equilibrates, this ratio should go to 1. This seems like a more objective metric than averaging over 200 seconds.

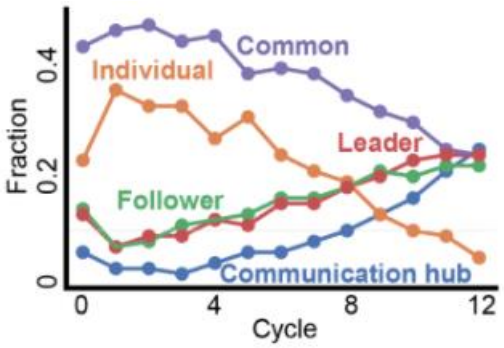
See above. We excluded the asynchronization measure and validated the adaptation rate with a parametric readout (new Fig. S1) so our readouts are now much less dependent on specific parameters.

Also, we replaced the measurements for heterogeneity (hub/individual ratio, bulk homogeneity) in communication with standardized network-science measurements (Estrada index and collective heterogeneity based on the cells' degree rank). See new Fig. 2D (posted in the response to reviewer #1).

2. Information transfer and the 5 classification types (Fig .2)

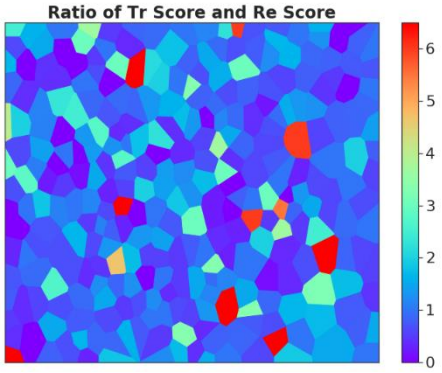
*** It would help to see a fraction plot (or pie chart) breaking down the fractions of the population that belong to each of the 5 defined categories**

This information (fraction of roles in the population over time) is shown in Fig. 5B, see below.

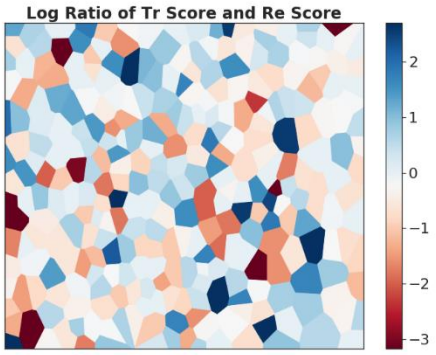


*** Fig. 2C is a bit confusing . It's hard to visually map the two images to look at the relative ratios of transmission/receiver. Can a ratio be shown over these two to give a better sense of where the 'hubs' are and how much a given cell can be both transmitter/receiver?**

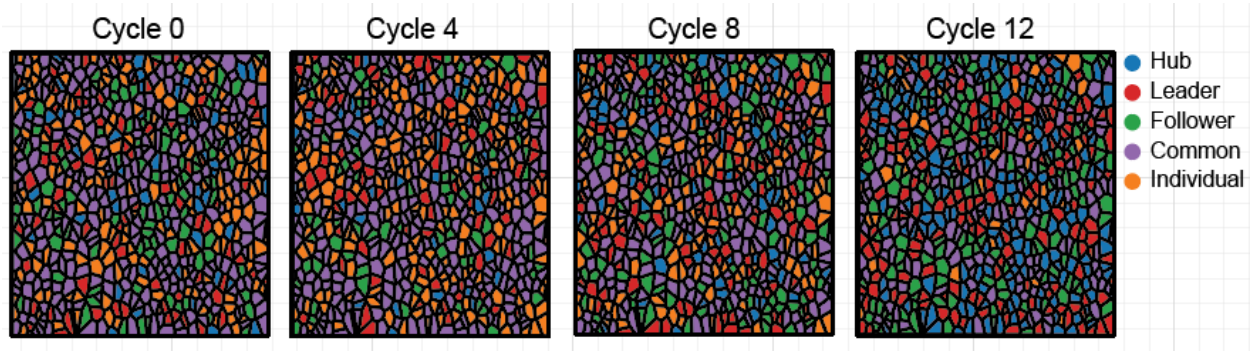
The ratio will not provide much information regarding the hubs because the ratio will be ~1 for "hubs". We provide here the ratio (transmission score/receiver score) for the reviewer's eyes below.



And the log (base 2) which is more informative:



We agree with the reviewer that better visualization of the receiver/transmission could be helpful. Note, that in the revised manuscript we introduce the transmission score/receiver score only in the "cycle" experiments, appearing after the "step" experiments, where we used the degree rank in our analyses. Thus, in the revised manuscript we provide a new supporting figure visualizing the color-coded roles over the cycles (Fig. 5B see below).

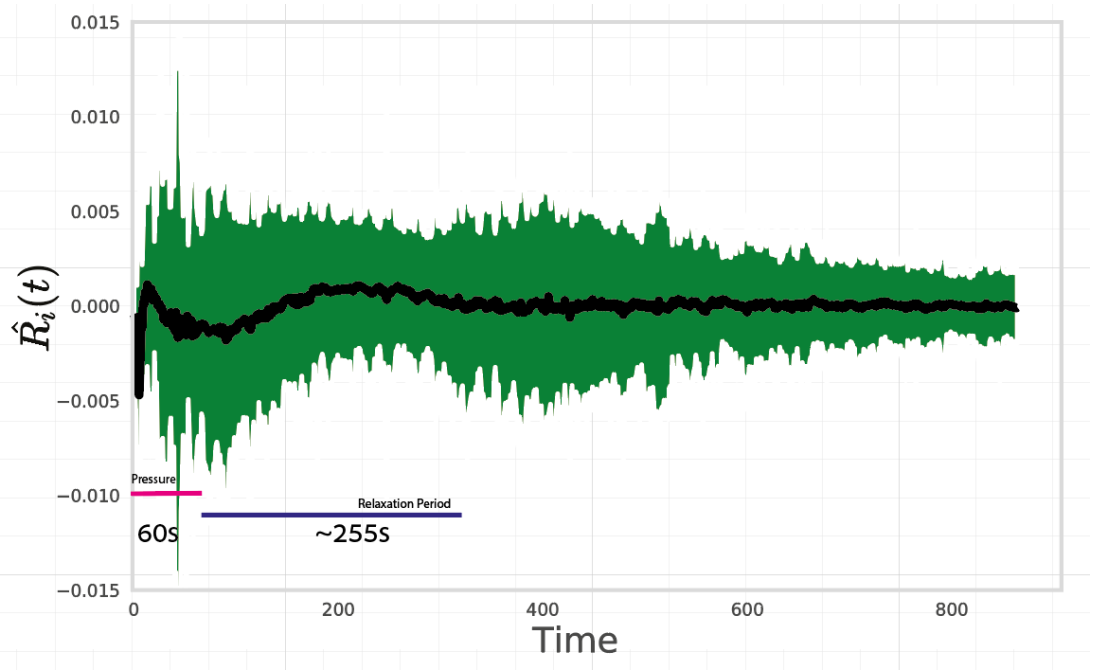


3. Periodic flow and memory (Figs. 3/4)

I have a specific biological and experimental concern here, and this is my primary concern with the paper and interpretations.

*** Show how the system responds to a single, 2 minute shear pulse with sufficient time to allow equilibration (plot standard deviation and R' for a single cycle with an additional 2-3 minutes after the shear is stopped). This will reveal the equilibration dynamics, as well as the general system response for a short pulse.**

We performed an experiment where we applied shear stress (2 Pa) for 60 seconds with sufficient time for relaxation. See below: mean calcium dynamics (black) and the standard deviation (green). As detailed above, the collective relaxation time is ~285 seconds, with single cell relaxation time ranging around ~70-110 seconds from a clear increase in brightness intensity near the time that the first pressure wave arrives (peak calcium activity).



*** Include control data prior to stimulation: As before, go back at least 1 full period (2 min+) for the plot above and the plots already shown.**

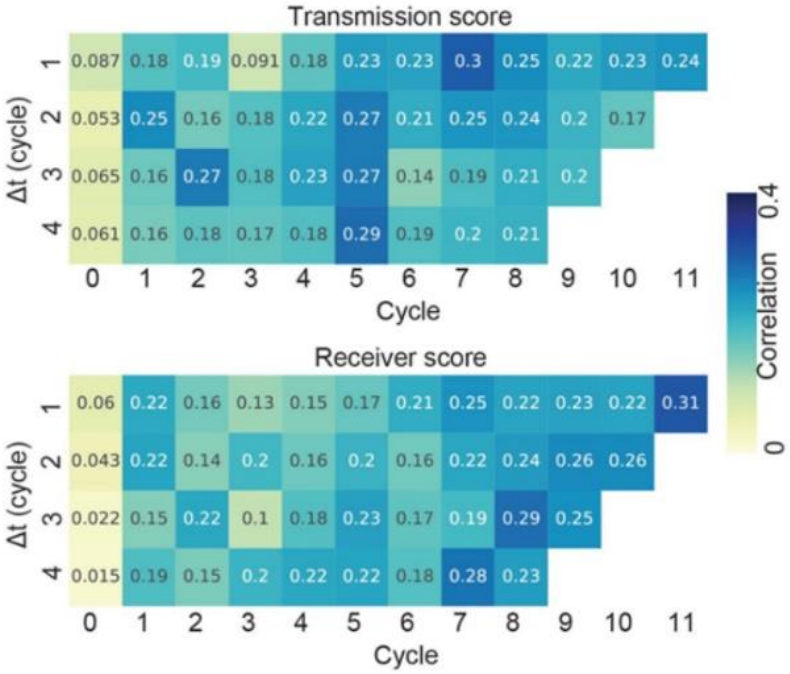
We showed above control data prior to the stimulation. As we argued before, the cells do not recover to their previous basal level intensity in the time-scales of our experiment. However, this does not prevent the cells to respond to a second pulse of stimuli. This is one reason why we used the derivative and not the raw intensity for our analyses.

*** Assumption of independence: The authors state that each cycle was treated as independent, and that each cycle was run over 2 minutes. However, when I look at the midline axis it looks to me like the R' signal never actually reaches equilibrium between cycles, especially early on. To me, this implies that the cells are still recovering from the Nth cycle when the N+1 cycle starts. While some of the cycles look like they may cross zero before the next begins, it's such a fine distinction that it still doesn't seem like the cells have had a chance to recover, meaning their behavior in the N+1 cycle may be biased by their behavior in the Nth cycle. This seems to matter because the analysis of cellular memory and how cells maintain or transition across roles depended on the assumption of independence across cycles. While the effect of tightening standard deviation is clearly real and meaningful regardless of independence, it's not clear to me how this relates to the concept of 'memory' especially with respect to questions of if a given cell has a preferred behavior. Would a cell that is a strong transmitter in cycle 1 still be such in cycle 2 if it were allowed to fully relax between cycles? Again, I may be misinterpreting this, but it seems like the experiment would need to be re-run with sufficient rest time between each cycle if independence is a necessary condition.**

As the reviewer note, the calcium dynamics never reach equilibrium between cycles because the shear stress is periodic and continuous. This means that the cells are not relaxing rather continuously synchronizing their calcium dynamics to the changing mechanical stimuli.

Regarding the reviewer concerns on our interpretation of "memory".

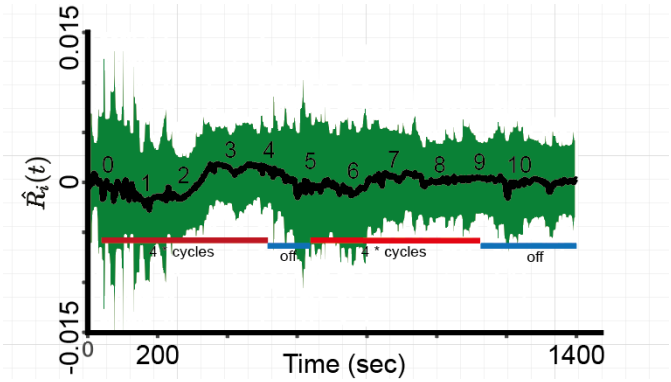
First, we found that the memory timescale is longer than a single cycle. There was no obvious decay in memory (correlation) up to four cycles apart, well beyond the timescale required for the multicellular system to relax (Fig. S8B, see below).



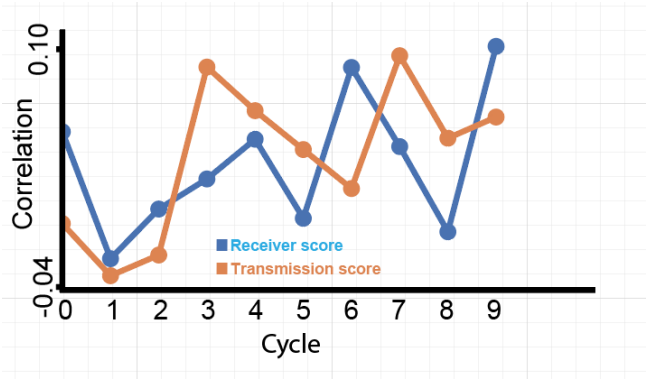
Second, estimating the single cells' relaxation time with pulse-release of shear stress showed that it takes on average ~70 seconds for a cell to return to basal activity from the moment that cell responded to the end of response. Together, we conclude that relaxation time is shorter than the time span between stress pulses, which is shorter than the memory decay time that expanded beyond 4 cycles that last ~480 seconds.

Third, we followed the reviewer's suggestion and performed an experiment with rest time between the applications of shear-stress periodic cycles. This was a tricky experiment. Even though in our single replicate the system did not synchronize well, we could still measure positive correlations between cycles, and specifically before and after the idle (i.e., pause in shear stress) cycle providing further evidence that the memory is a cell property that is independent across cycles. See below.

This is the calcium dynamics over time. The pause is shear stress was during cycle #4.



This is the memory (i.e., single cell correlation in receiver/transmitter score) with time delay of two cycles. Note the (low but) positive correlation that is increased over cycles. Specifically, the positive correlation at cycle #3 (x-axis, between cycle 3 and 5) including the idle cycle.



Because these memory measurements were less convincing than other "cycle" experiments (see new Table S2) we decided not to include these results in our revised manuscript. We are willing to replicate this experiment and include it in the manuscript if the reviewer and editor believe this is necessary.

4. Topological distance discussion (Fig. 5)

I'd like to see more discussion of what these plots mean, what spatial permutation entails, and how to interpret the data in a broader context.

We elaborated on these results in the Results and in the Discussion.

From Results - "Our data suggests multicellular synchronization is associated with various single cell properties such as communication state and memory. This led us to the hypothesis that the synchronization process is driven by effectively propagating information from the local scale (between single cells), to the global (collective) scale. To test this hypothesis, we measured to what extent local cell properties explained the information flow in the multicellular network. First, we computed the neighboring pair cross correlation coefficients for direct observations and

spatially permuted data. We found that the spatial permutation always decreased the cross correlation, therefore cross correlation was maintained as a local cell property throughout the experiment even in the presence of common external stimuli (Fig. 6A, upper-left inset). Intriguingly, spatial permutation decreased the GC edge probability in early cycles but increased the edge probabilities in later cycles (Fig. 6A main panel and the lower-right inset). These results indicate that once neighboring cells reach sufficient synchronization their ability to influence each other is less effective than cell pairs far apart. We validated these observations more systematically by correlating the topological distance between pairs of cells to their GC edge probability (Methods). This analysis established that at the onset of the experiment, the information flow is dominated by local cell-cell interactions and is gradually transitioning to the global scale as the multicellular network synchronizes (Fig. 6B)."

From Discussion – "Our study reveals a self-organized multicellular network that supports information flow from local to global scales. Such information may be carried by two main signaling mechanisms, juxtacrine (contact-dependent) and autocrine (secreted-dependent) {Fancher, 2017 #1278}. A juxtacrine channel allows a cell to establish conversation with its (physically touching) immediate neighbors without interference from extracellular space. For HUVEC cells such communication can be realized by gap junctions {Okamoto, 2017 #1279}. On the other hand, an autocrine channel allows a cell to broadcast its information through diffusive messengers in the extracellular space. For HUVEC cells stress-triggered ATP release and ATP-induced calcium dynamics constitute an autocrine pathway {Yamamoto, 2011 #1283}. While both mechanisms could contribute to the information flow within the multicellular network, we suggest gap-junction and contact-dependent signaling as the dominant mechanism (Fig. 3B versus Fig. 3C or Fig. S5). While a recent study suggested that positive feedback of a diffusive signaling mechanism can drive accelerated, long-range information transmission {Dieterle, 2019 #1287}, the external flow in our system is likely to rapidly dilute the diffusive messenger {Gregor, 2010 #1284}. The contact-dependent information flow hypothesis is also supported by our previous studies where we demonstrated that blocking gap junctions, or inserting weakly communicating cells impaired the information flow {Sun, 2012 #255;Potter, 2016 #760}."

We look forward to your evaluation of this submission.

Best regards,

Assaf Zaritsky, Ph.D.

Ben-Gurion University of the Negev, Be'er Sheva, Israel. Email: assafza@bgu.ac.il, lab website: https://www.assafzaritsky.com/

אוניברסיטת בן-גוריון בנגב
Ben-Gurion University of the Negev
Faculty of Engineering Sciences
Software and Information Systems Engineering

Reviewer comments:

Reviewer #2: The authors have addressed all the issues I had. The paper is awesome and should be published in Cell Systems.

Reviewer #3:

Reviewer #3: First, my sincere apologies to the authors for the long delay. My family experienced a major medical crisis with ongoing complications. My intent in agreeing to re-review had been to spare the authors adding an additional and unfamiliar reviewer, but I fear that I was forced to delay longer than anticipated. I appreciate the authors' patience and understand that they also experienced a difficult past year.

That said, I appreciate the extensive reviews and thoughtful revisions. I still have concerns that should be clarified, but do not feel the need to review this again.

The #s and headers below correspond to those from my original review.

1.1: I still don't understand 1C or what it means to say 'dashed lines are individual calcium' responses when I don't see dashed lines here and I'm not sure of the connection between 1C and 1D.

We are sorry for this mistake. We corrected “dashed lines” to “colored lines” and “solid line” to “black line” in the legend of Fig. 1C.

Again, minor concern but why not just show the pure 0.2 Pa signal in 1C with AU and then move to R' in 1D?

Fig. 1C shows that “the magnitude of the cell’s calcium signal correlated with the magnitude of the applied flow shear stress", this is a control experiment before moving to our “standard” setting of a “step”-like shear stress in Fig. 1D.

Regardless, I think it's essential to show when stimulation turns on/off in the plots. This was not clear originally in the plots, and I think something like the green lines in the rebuttal would help in Fig. 1C.

This information is provided in Fig. 1C (dashed horizontal lines), but was not properly explained in the figure legend. We now included “Dashed horizontal lines indicate the time interval of 0, 0.2, 0.3 Pa shear stress correspondingly.” In the legend.

In the rest of the “step” experiments results t = 0 is the time when shear stress is first applied. This now explicitly stated in the figure legends.

1.2: I understand that the authors feel that their visualization choice is "easier to grasp by the reader", but as a representative reader I have to disagree and maintain my request. I do appreciate that the authors showed a representative example of R(t) in the rebuttal, but this plot doesn't really look like the inset in 1D, so that confuses me--why doesn't it normalize to something as striking as shown in Fig. 1D's upper inset? It's fine if this is batch-to-batch variability, but I think it's important to at a minimum include these data in the supplement and explain the difference between these and 1D, upper inset.

We moved this new figure from the rebuttal to the SI (Fig. S13). In this experiment the imaging starts before the mechanical stimuli onset and then stops (130s idle, then 60s shear flow, and then idle again). This is a very different experimental setting from the one shown in Fig. 1D where the flow onset and imaging onset coincide. Thus the difference in the time patterns between the experiments is not surprising. The same experiment was used to assess single cell relaxation time (see next point).

1.3: I appreciate the rephrasing and clarification of the metrics.

*** Response to request for comparison to control data and if cells actually return back to control fluorescence.**
Here, the rebuttal explains how single cells relax at 70-110 seconds, while groups relax collectively at 285 seconds. The fact that cells relax to greater than their initial control value is now shown in 1D, inset and I appreciate that (and find it interesting).

I do think it's important that the authors discuss where their estimates of relaxation time come from (how did they determine steady state, etc.) and include the relaxation data as a supplemental figure as it's relevant to the other figures and interpretations.

These data is also shown in Fig. S13. The figure legend now includes an explanation how the relaxation time was estimated: “For each cell, the calcium peak near the time to the pulse of shear stress application and the time where the cell settled (or “relaxed”) were manually annotated”.

2. Information transfer and the classification types
2.1/2.2 -- I appreciate the visualization discussion and find the Fig. 5B update along with 5C to be a helpful visualization.

3. Periodic flow and memory (Figs. 3/4)
I appreciate the authors undertaking the increased rest-time experiments and clarify memory. However, there are a few things I still do not understand here.

First, I think the gap junction data are quite striking and very compelling. This is a strong datapoint supporting some form of coupling.

*** Most importantly, the authors re-iterate in the text that they are "treating each round as an independent cycle, and comparing single cell responses across cycles..." My original**

concern was that it seems like cycles can only be independent if cells are allowed to reach a steady state between cycles, otherwise it is unavoidable that the cell will still be experiencing effects from the Nth cycle when the N+1th cycle begins. The authors appear to agree that calcium dynamics never reach steady state between cycles in the rebuttal, but then add that "cells are not relaxing rather continuously synchronizing their calcium dynamics to changing mechanical stimuli." I feel like this remains a major point of rhetorical confusion. If the cycles must be understood as truly independent, as is written in the main text (see above), then this interpretation that the cells are continuously synchronizing to a changing stimulus seems problematic for the assumption of independent cycles. I think confusion comes from the use of the words 'memory' and 'independence'. If you provide a continuously varying signal to a system and the system becomes 'better' at anticipating that signal and the standard deviation tightens, is that a sign of 'memory' or a sign of 'entrainment' (it's an excitable system that never relaxes between cycles). I do understand the argument that the cyclic stimulation data reflects a system that is better able to follow a continuous stimulus. I am probably missing something, but as not every reader will be as well versed in the correlation analyses used here better discussion in the main text is essential to clarify this.

As the reviewer nicely articulated, this is a 'rhetorical confusion', rather than a scientific one.

Although both words have relevance here, we think the word ‘memory’ is more suitable than ‘entrainment’ in this context because cellular memory refers to cells tending to retain their functional role (leader/follower/hub) across cycles. Entrainment, to our understanding, refers to the phase locking of a cell to its neighbors or to the external stimuli - i.e., the system becomes 'better' at anticipating a continuously varying signal leading to tightening the standard deviation.
Thus, entrainment is more related to the synchronization aspect, rather than the cell-to-cell heterogeneity. We believe that the word memory is more suitable because we show multiple evidences that synchronization is a result of communication, and not because independent adaptation of the single cells to the external stimuli. In fact, to some extent memory and entrainment have opposite effects: while the former preserves cell-to-cell heterogeneity, the latter emphasizes the process of reaching uniformity.

We would also like to argue that it is not necessary to build memory after complete relaxation. On the contrary, from behavior memory, to genetic memory, to non-living systems such as thermoplastics, memory is associated with long-term internal modifications.

We are treating each stimuli cycle as independent events. We hypothesize that the cells are accumulating molecular modifications to stabilize their functional role against random fluctuation, which we refer to as memory. These molecular modifications may have multiple manifestations, including the calcium level not completely returning to the basal level. So we think it is not contradictory to have cellular memory when cells are not fully relaxed from the previous cycle. To make our point clear, we have included the following two sentences to the main text: (1) in the Methods: “each stimulus cycle was considered as an independent event, although the cells’ calcium dynamics ($(R_i)\tilde{}(t)$) never reached equilibrium between cycles because the shear stress was periodic and continuous”, (2) in the Discussion: “While admittedly

further investigations are required to reveal the molecular mechanisms of the cellular memory, we suspect slowed gap junction turnover, as well as the continuously perturbed calcium dynamics from fully relaxation, may contributed to the reinforced functional role of cells."

*** Regarding the experimental data shown with relaxation between cycles, I appreciate that the authors attempted this experiment and were willing to show the raw data but I am confused about this experiment was conducted. I had asked for the experiment to be run using a single pulse followed by sufficient time for relaxation followed by a new pulse, etc. Here, 4 cycles were run serially followed by a 2 minute gap (I think). In no cycle was there any response resembles that shown in Fig. 3. The lack of reproducibility is surprising—is this typical or expected and would it interfere with someone following up on this work? This also seems to have been the case in some of the other R' plots shown in the rebuttal document where they do not replicate the strength of the phenomena shown in the actual figures. Again, I appreciate the authors' transparency, but feel like this should be discussed if the system is particularly susceptible to something like passage #, time post-thaw, or individual thaws, for example.**

The magnitudes of the cell responses vary significantly across experiments. HUVEC cells are primary cells where each batch could come from different donors. Like other endothelial cells, HUVEC cells may have a mixture of quiescent and sprouting phenotypes. Although we have tried our best to control the culture condition, it is conceivable that variations are intrinsic to the system. Therefore, our analysis has been devised to a large extent overcoming such variations. For instance, by using the derivative of intensity that is independent of the basal level, by using the z-score to normalize different experiments, and by using Granger inference that is independent of scales. Indeed, in the manuscript we show one of the experiments where these results were more prominent, but it is important to explicitly state that we successfully reproduced all the key findings, and reported all of the replicate experiments' results in Table S2.

Specifically, the tricky experiment in the rebuttal, rest time between the applications of shear-stress periodic cycles, did not synchronize well. However, even with these difficulties, we still measured positive correlations between cycles, and specifically before and after the idle (i.e., pause in shear stress) cycle providing further evidence that the memory is a cell property that is independent across cycles. We now include these results in the SI (Fig. S15).

Minor:

*** Fig. 3 repeats several times that there were 13 cycles performed, but I see 15 peaks in the plots. This is relevant because I'm having a hard time understanding where the cycles stop and start and was trying to calculate cycle time. This point needs clarification. Again, including fiducial markers in all plots for when stimulation stops and starts would be a service to the reader.**

We included visual markers to clarify the time of each cycle. We limited the number of cycles per experiment to 13 and excluded further cycles, when such existed (as in the figure). This is now

אוניברסיטת בן-גוריון בנגב
Ben-Gurion University of the Negev
Faculty of Engineering Sciences
Software and Information Systems Engineering

explicitly stated in the Methods: “We limited the number of cycles for analysis to 13 at most, because of gradual accumulation of stage drifting and photo-damaging effects”, and we excluded the extra cycles that were not analyzed in Fig. 3B.

*** I can't interpret Fig. S8B. I'm not really sure how to read the tables, and I don't see why a correlation score of 0.3 is considered a good sign of memory. It seems like a weak behavior at best here, but again this is not my specialty. I also don't understand Table S2 at all, both due to the visual presentation (across many pages) and the vast amount of data grouped together. Again, I am likely missing something but I don't see a smoking gun for memory here.**

Correlation of 0.3 is well beyond random implying that the cells have a memory mechanism (along with other mechanisms that affect the single cell communication properties). This correlation is increasing across cycles and these results are replicated across multiple experiments (Table S2) implying that it is a real phenomenon. In Fig. S8 we show that this correlation holds for way longer than a single cycle. There was no obvious decay in the correlation up to four cycles apart, which is beyond the timescale required for the multicellular system to relax. We clarified in the figure legend that the x-axis are the cycles and the y-axis the number of cycles considered when calculating the correlation.

Table S2 - this is just a formatting issue. The raw file is excel and should be formatted properly in the published version.

*** I think it would definitely help to provide additional context for all assumed-commonplace metrics like correlation scores. For some of us, a correlation of 0.3 is problematic, but I gather that is a good thing here, for example. There is a lot of discipline specific jargon and assumptions here and stepping back in the writing every now and then to add a sentence or two in a few key places might help improve accessibility.**

We included more explanations throughout the manuscript to make it easier to follow.

We look forward to your final approval and publication of this submission.

Best regards,

Assaf Zaritsky, Ph.D.

Ben-Gurion University of the Negev, Be’er Sheva, Israel. Email: assafza@bgu.ac.il, lab website: https://www.assafzaritsky.com/