How to design effective figures for scientific* articles

Presented by Stacey Chin, Thomas Dursch, and Matthew Pavlovich

29 October 2020

About the speakers

Tom Dursch, PhD
Trends in Chemistry
tdursch@cell.com
@TrendsChemistry

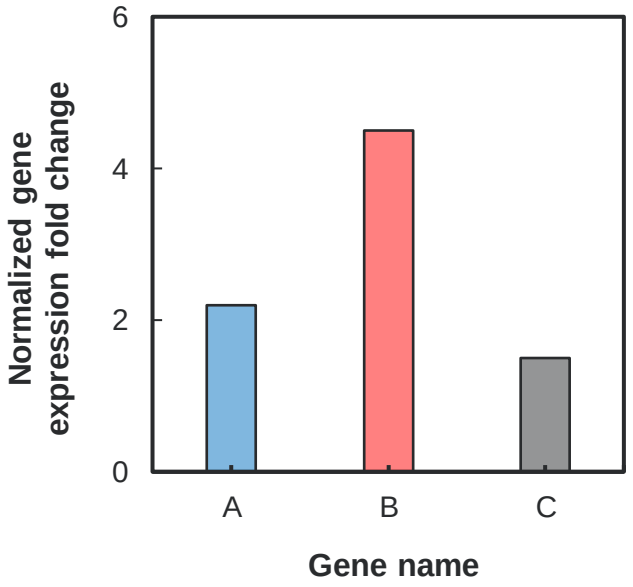
Stacey Chin, PhD
Matter
s.chin.1@cell.com
@stacey_chin

Matt Pavlovich, PhD
Trends in Biotechnology
mpavlovich@cell.com
@TrendsinBiotech

The purpose of scientific figures

- Scientific figures are designed to illustrate, introduce, and/or emphasize ideas/data that are otherwise difficult or lengthy to explain via text
- Two key considerations when designing a scientific figure: (1) identify the message and (2) consider the audience

	Time post-infection (hr)			
Gene Name	1	2	4	12
A	1.0	2.2	1.0	0.15
B	1.0	4.5	0.95	0.05
C	1.0	1.5	2.0	2.0



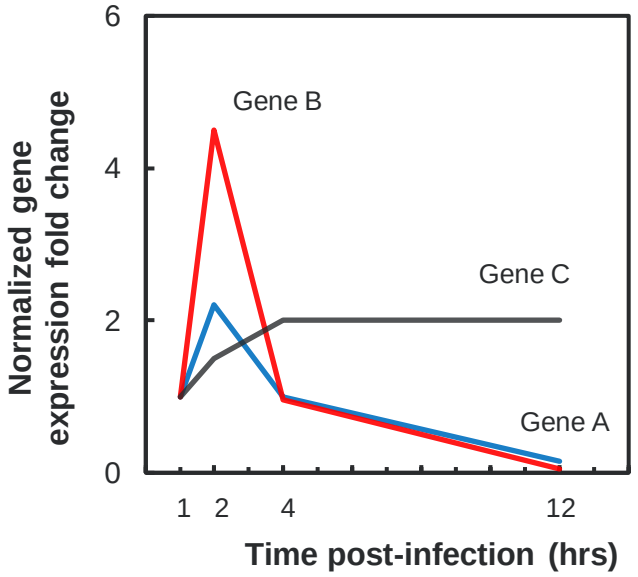
https://mitcommlab.mit.edu/broad/commkit/figure-design/

https://doi.org/10.1371/journal.pcbi.1003833

The purpose of scientific figures

- Scientific figures are designed to illustrate, introduce, and/or emphasize ideas/data that are otherwise difficult or lengthy to explain via text
- Two key considerations when designing a scientific figure: (1) identify the message and (2) consider the audience

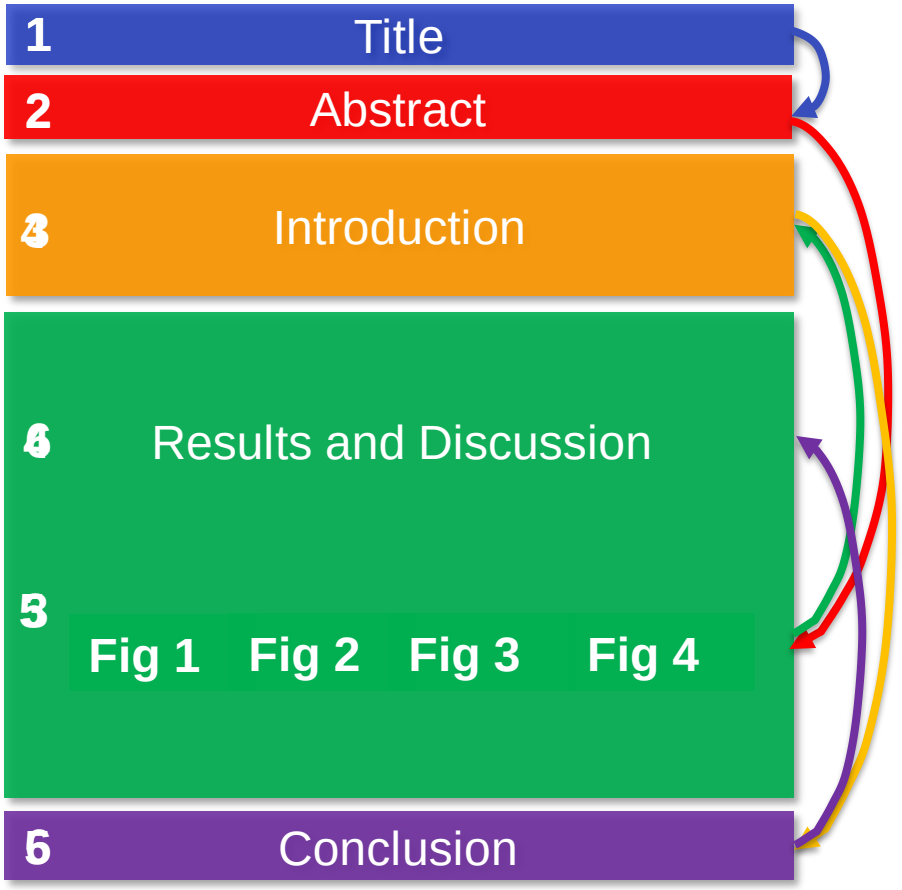
	Time post-infection (hr)			
Gene Name	1	2	4	12
A	1.0	2.2	1.0	0.15
B	1.0	4.5	0.95	0.05
C	1.0	1.5	2.0	2.0



https://mitcommlab.mit.edu/broad/commkit/figure-design/

https://doi.org/10.1371/journal.pcbi.1003833

The importance of scientific figures



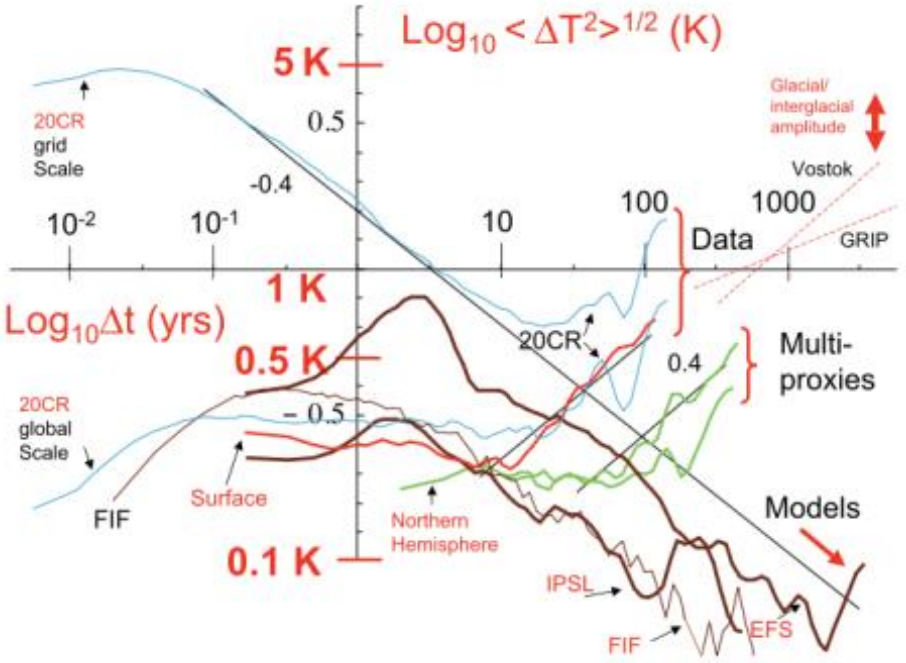
In an era of information overload, first impressions are critical!

General guidelines for figure design

- Identifying both the purpose of and audience for the figure allows one to best design an illustration that expresses the intended message

- There are many types of scientific visualizations, but below are some common items to consider:

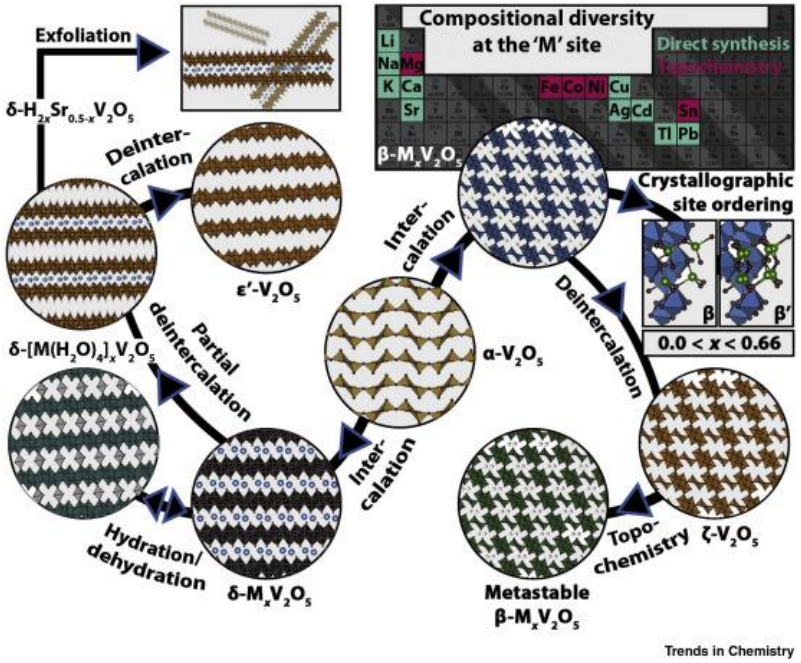
  - **Information density**



https://betterfigures.org/2012/11/28/too-much-information/

General guidelines for figure design

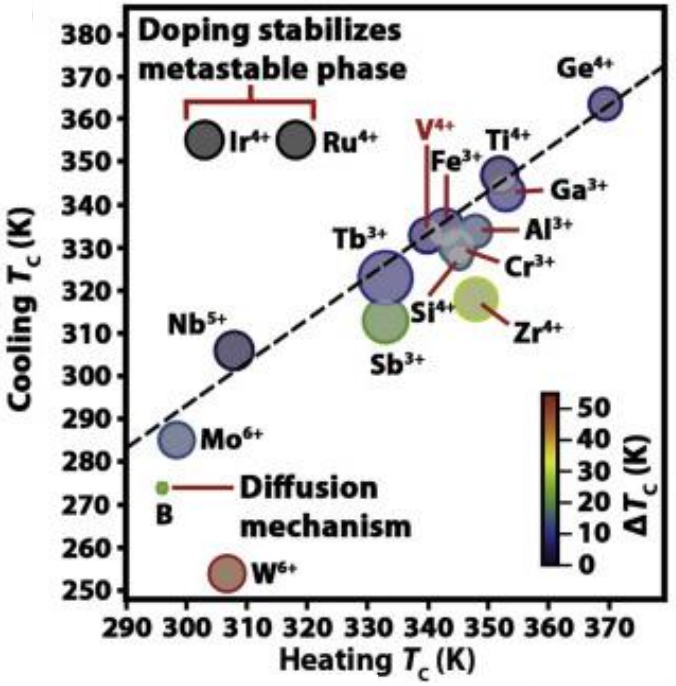
- Identifying both the purpose of and audience for the figure allows one to best design an illustration that expresses the intended message
- There are many types of scientific visualizations, but below are some common items to consider:
  - Information density
  - **Information flow**



https://doi.org/10.1016/j.trechm.2019.07.005

General guidelines for figure design

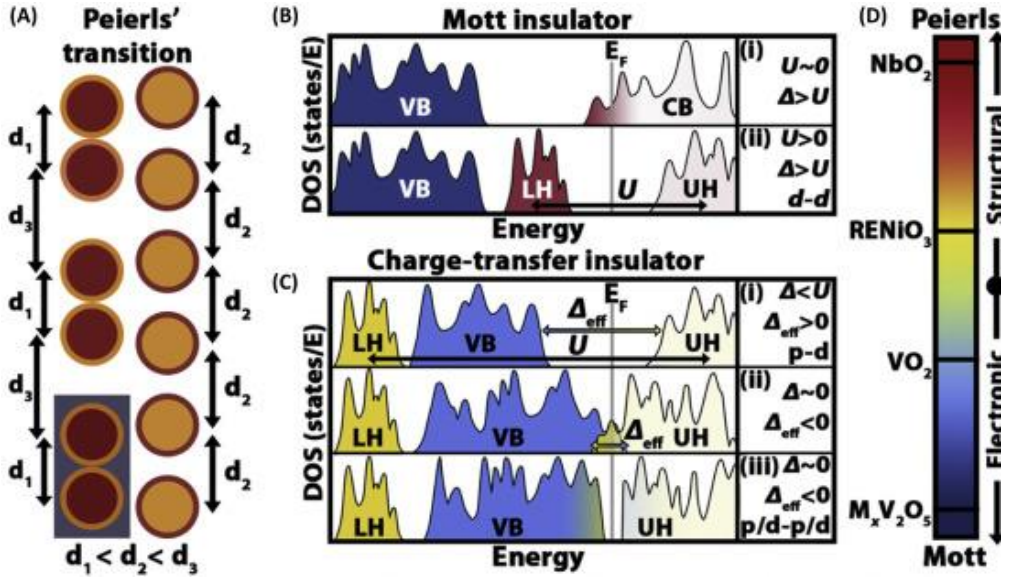
- Identifying both the purpose of and audience for the figure allows one to best design an illustration that expresses the intended message
- There are many types of scientific visualizations, but below are some common items to consider:
  - Information density
  - Information flow
  - **Scientific rigor**



https://doi.org/10.1016/j.trechm.2019.07.005

General guidelines for figure design

- Identifying both the purpose of and audience for the figure allows one to best design an illustration that expresses the intended message
- There are many types of scientific visualizations, but below are some common items to consider:
  - Information density
  - Information flow
  - Scientific rigor
  - **Figure aesthetics**



https://doi.org/10.1016/j.trechm.2019.07.005

General guidelines for figure design

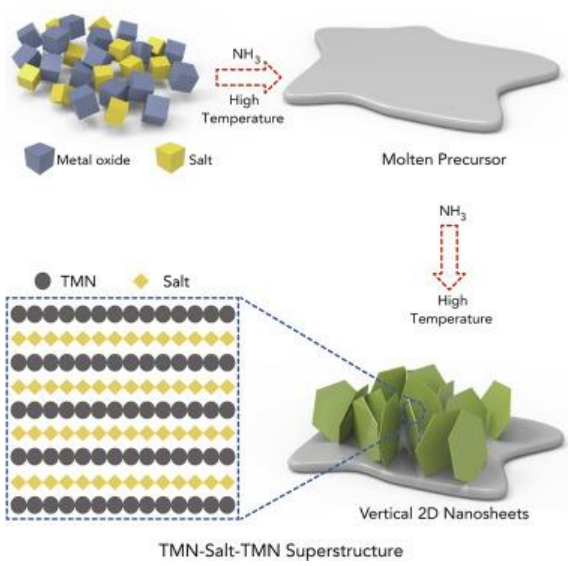
- Identifying both the purpose of and audience for the figure allows one to best design an illustration that expresses the intended message
- There are many types of scientific visualizations, but below are some common items to consider:
  - Information density
  - Information flow
  - Scientific rigor
  - Figure aesthetics
  - **Figure description/caption**

Figure 3. Metal–Insulator Transitions (MITs) in Vanadium Dioxide (VO_2). (A) Thermally induced MIT of VO_2 nanowires as measured by differential scanning calorimetry (DSC). The hysteresis becomes more pronounced with increasing ramp rate. (B) VO_2 can be doped with homovalent (e.g., Ti^{4+}) or aliovalent (e.g., W^{6+}) cations to alter the heating and cooling critical temperature (T_C) of the MIT. Reports of T_C modulation on the incorporation of various dopants is summarized

Classes of scientific research figures

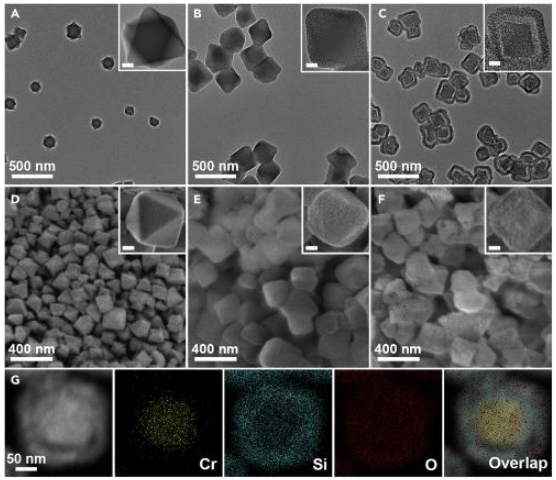
- What kind of information is your figure trying to convey?

Illustrative
"This is our plan"



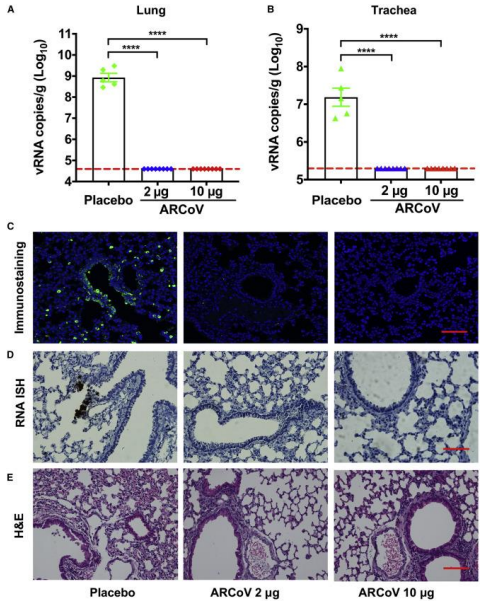
Jin, Huanyu, et al. *Chem* 6.9 (2020): 2382-2394.

Characterization
"This is what we made"



Bao, Shouxin, et al. *Matter* 3.2 (2020): 498-508.

Applications
"This is what it does"

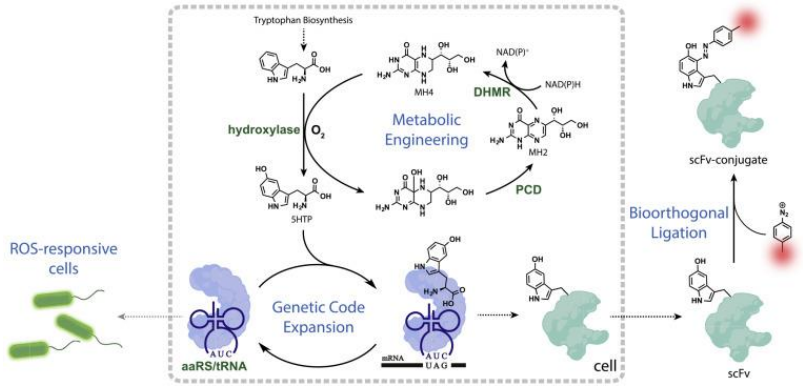


Zhang, Na-Na, et al. *Cell* 182.5 (2020): 1271-1283.

Illustrative figures

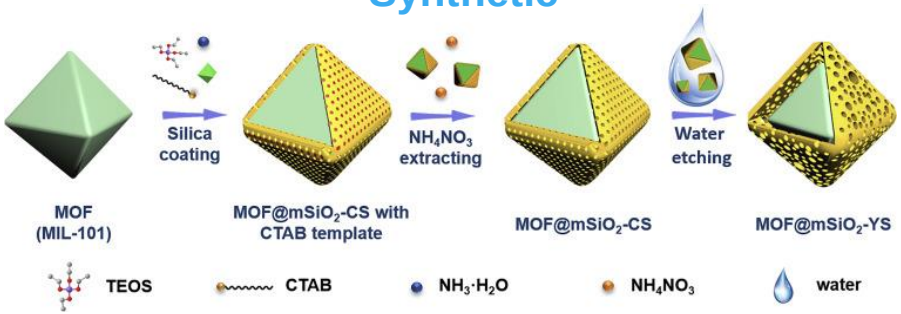
- A **high-level** look to give context to the data & improve reader understanding

Schematic



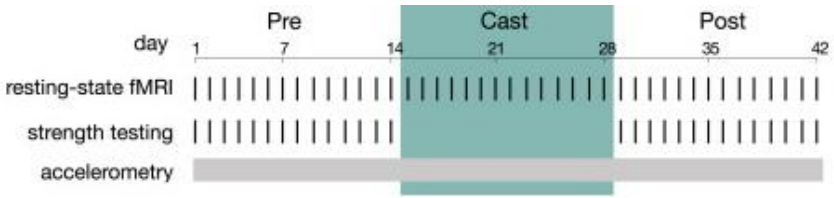
Chen, Yuda, et al. *Chem* 6.10 (2020): 2717-2727.

Synthetic



Bao, Shouxin, et al. *Matter* 3.2 (2020): 498-508.

Experimental Design

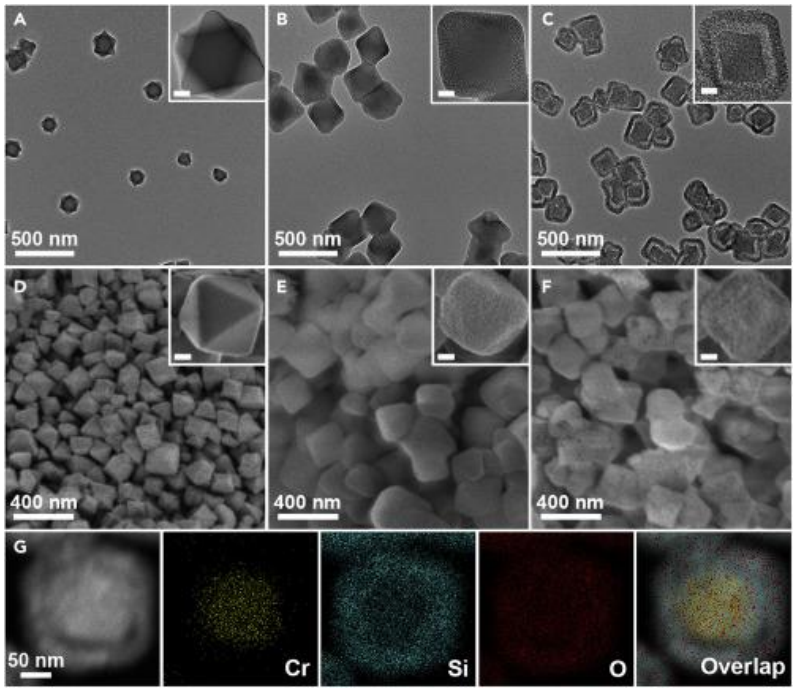


Newbold, Dillan J., et al. *Neuron* 107.3 (2020): 580-589.

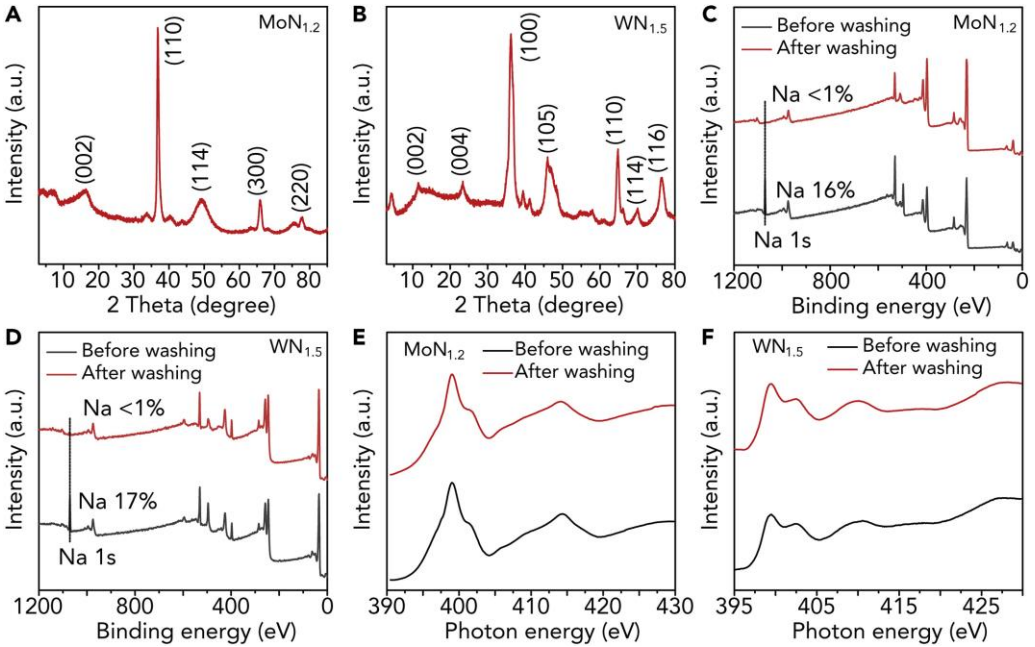
There is no such thing as being **too clear**!

Characterization figures

- An **in-depth** look into the system and how it functions
- Tend to be data-heavy (e.g., spectroscopy and microscopy)



Bao, Shouxin, et al. *Matter* 3.2 (2020): 498-508.



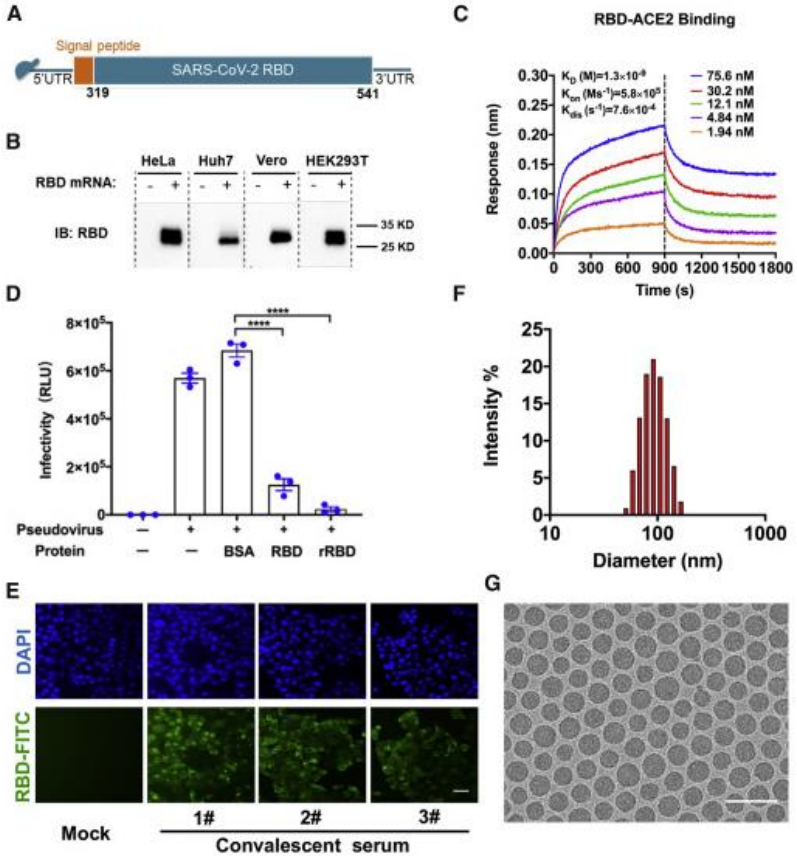
Jin, Huanyu, et al. *Chem* 6.9 (2020): 2382-2394.

The average number of data items and average number of panels per figure has **doubled** in the past two decades

Cordero, Radames JB, et al. *PLoS One* 11.6 (2016): e0156983.

Characterization figures

- An **in-depth** look into the system and how it functions
- Tend to be data-heavy (e.g., spectroscopy and microscopy)

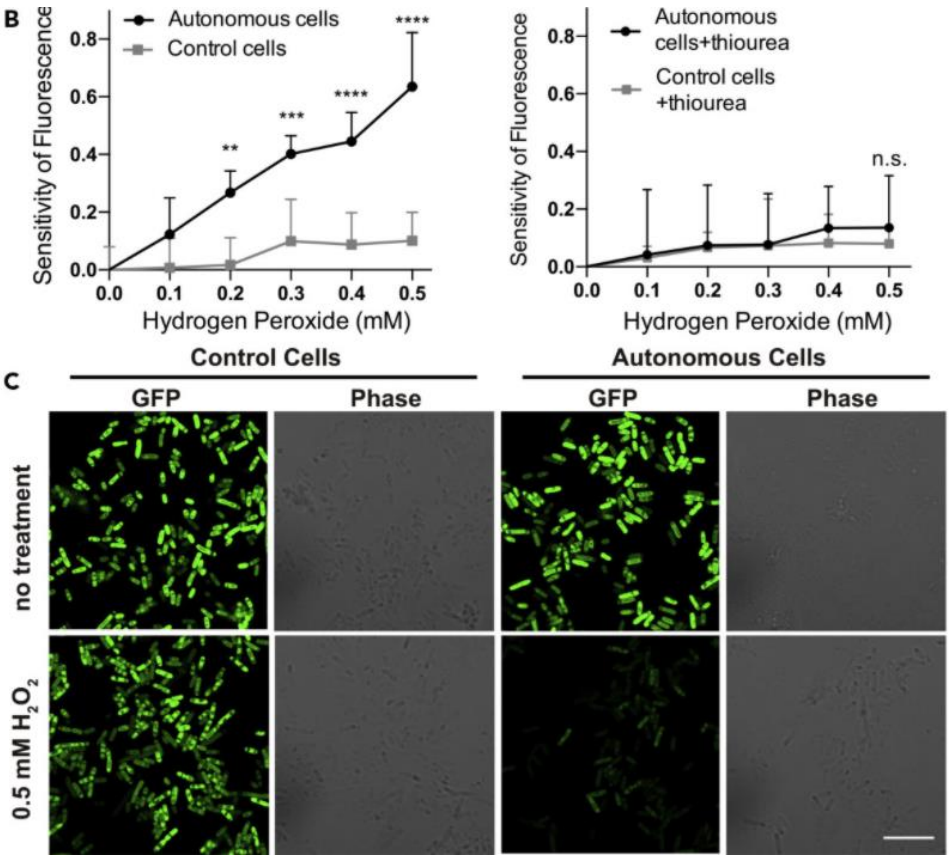


Zhang, Na-Na, et al. *Cell* 182.5 (2020): 1271-1283.

Be critical about what is **necessary** vs **nice to have**

Applications figures

- **Proof-of-concept** to show the potential of the system

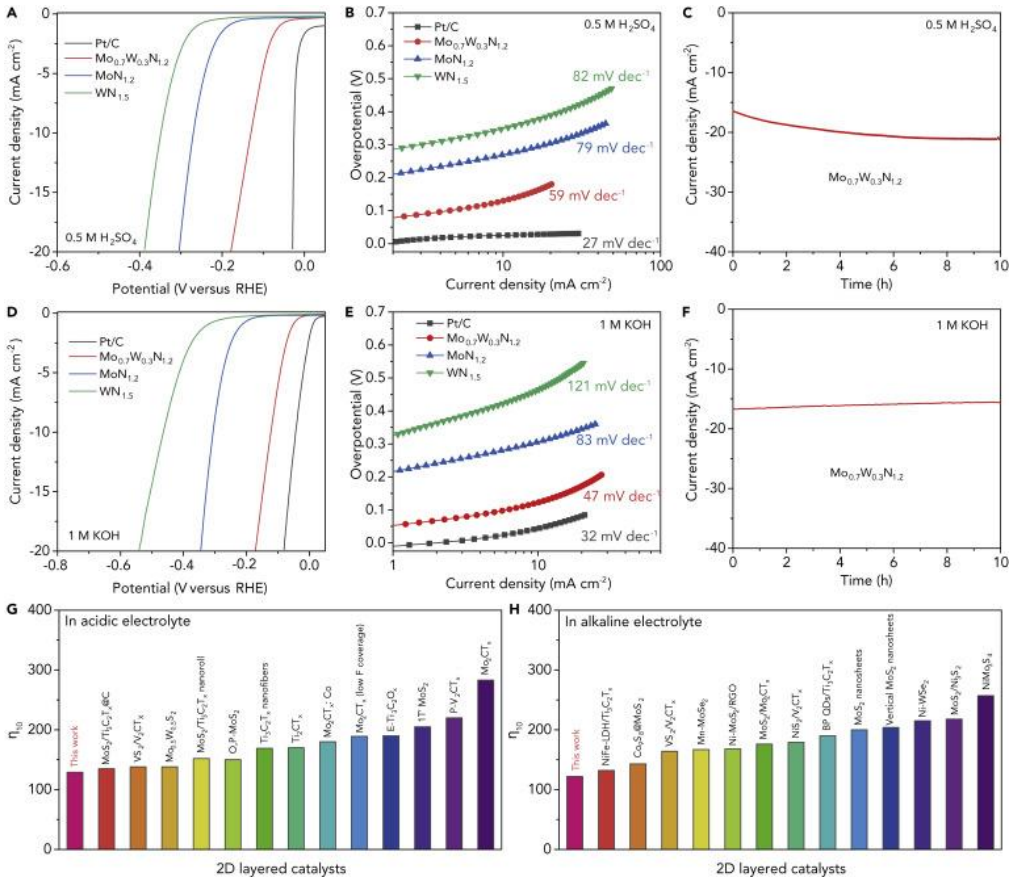


Chen, Yuda, et al. *Chem* 6.10 (2020): 2717-2727.

Qualitative vs. **Quantitative** data presentation

Applications figures

- **Proof-of-concept** to show the potential of the system

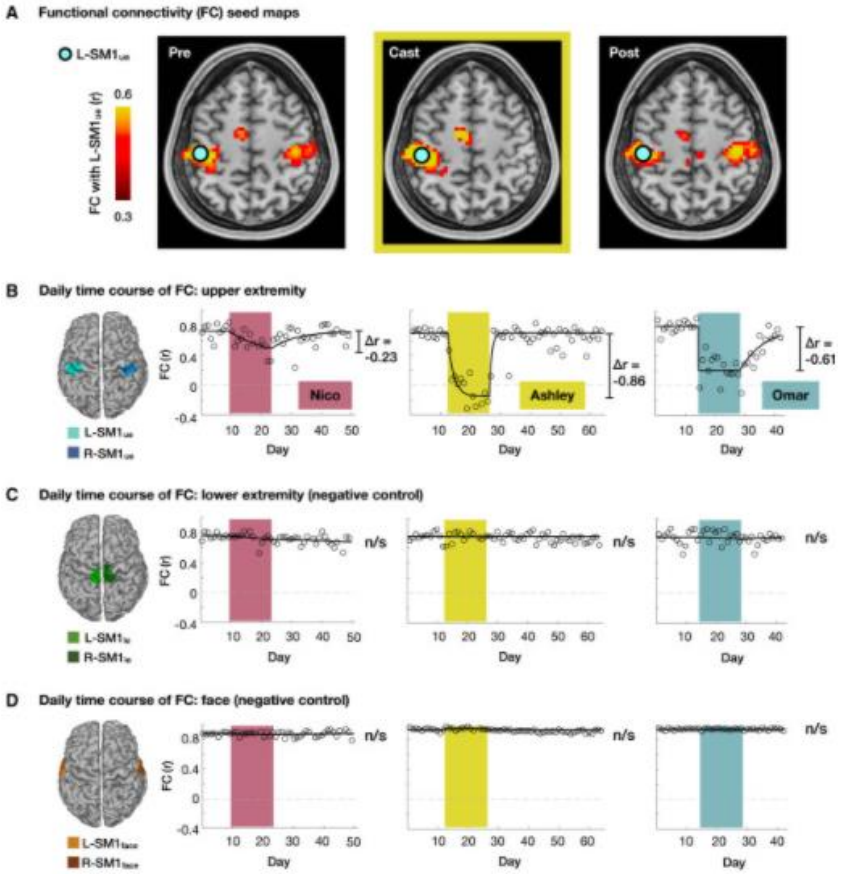


Jin, Huanyu, et al. *Chem* 6.9 (2020): 2382-2394.

Benchmark comparisons to other systems

Applications figures

- **Proof-of-concept** to show the potential of the system

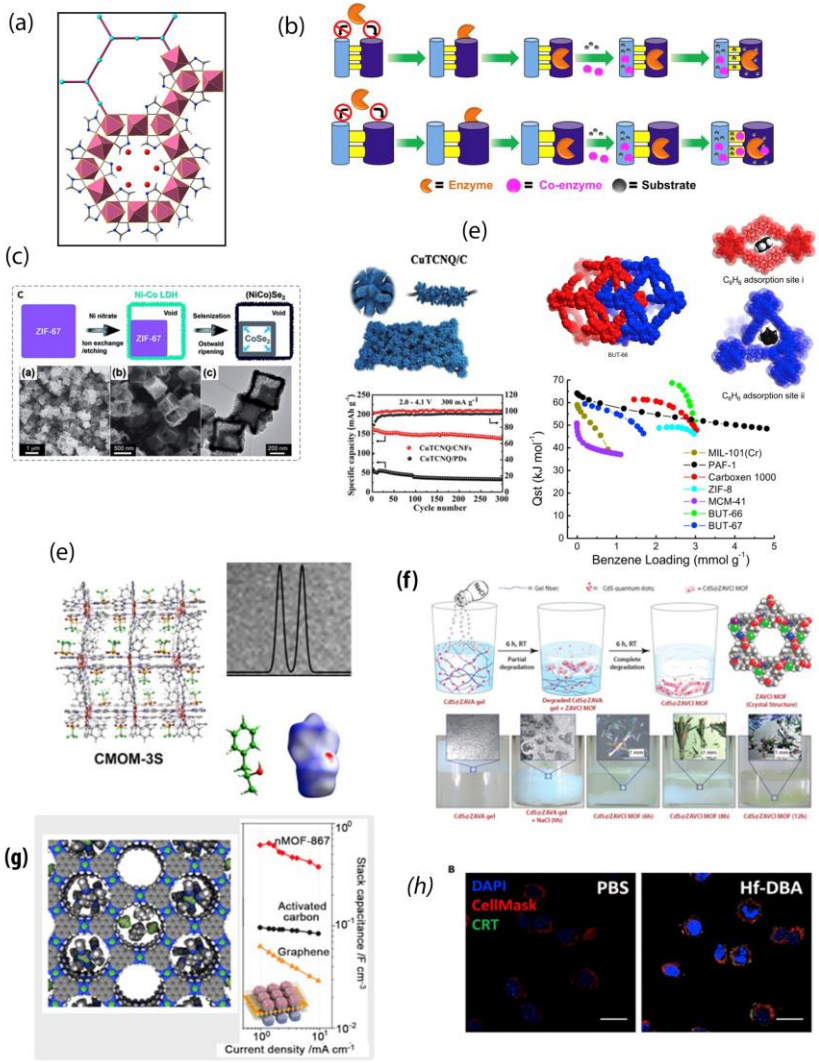


Newbold, Dillan J., et al. *Neuron* 107.3 (2020): 580-589.

What should the **main takeaway** from the paper be?

The dreaded Frankenfigure

Figure 1. Applications of MOFs.

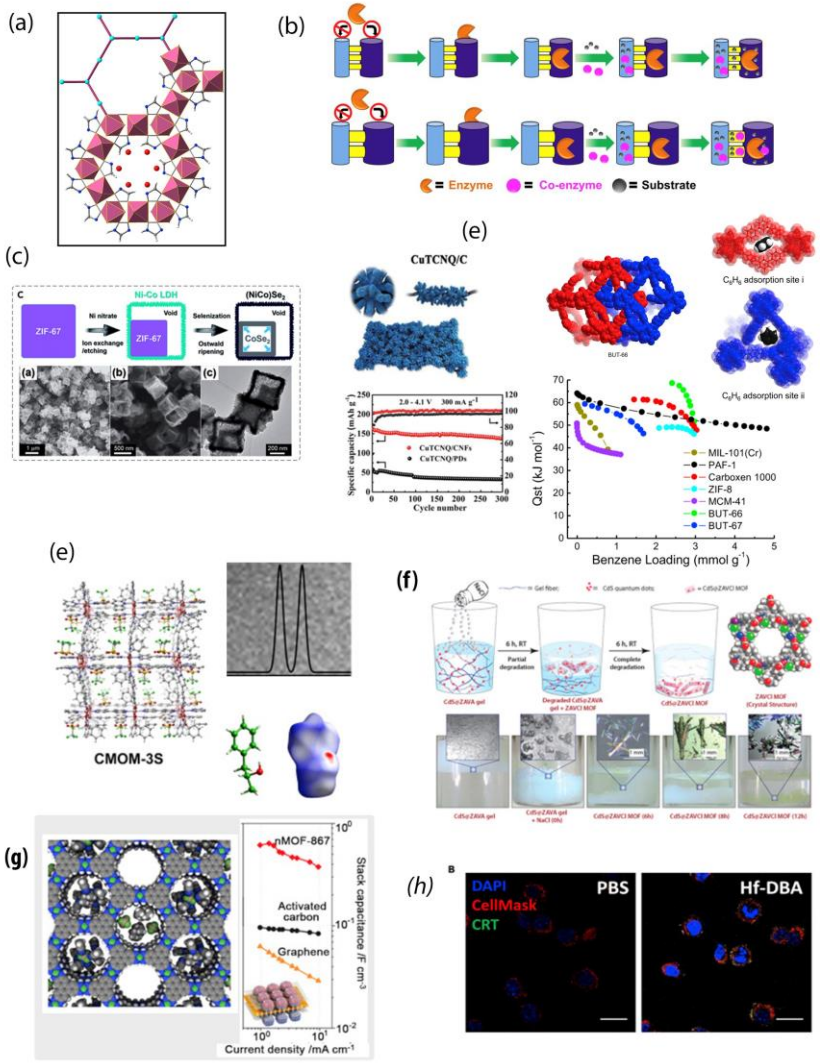


What is "wretched" about this?
What "horrible contrasts" can you find?

- No logical arrangement to the panels
- Unexplained acronyms
- Duplicated (c), missing (d), or inconsistent (h/B) panel labels
- Unlabeled chemical structures
- Unlabeled plots
- Inconsistent row/column layout
- Inconsistent level of detail in different panels
- Font styles, sizes, and backgrounds vary between panels
- Inconsistent panel framing
- Illegible image labels
- **No new insight beyond what the authors of the original studies already published**

The dreaded Frankenfigure

Figure 1. Applications of MOFs.

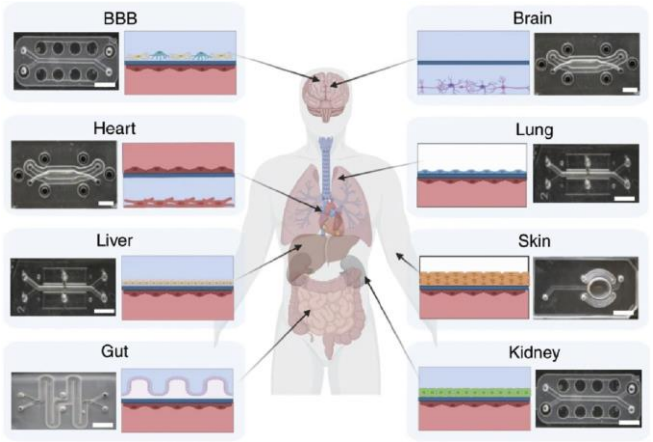


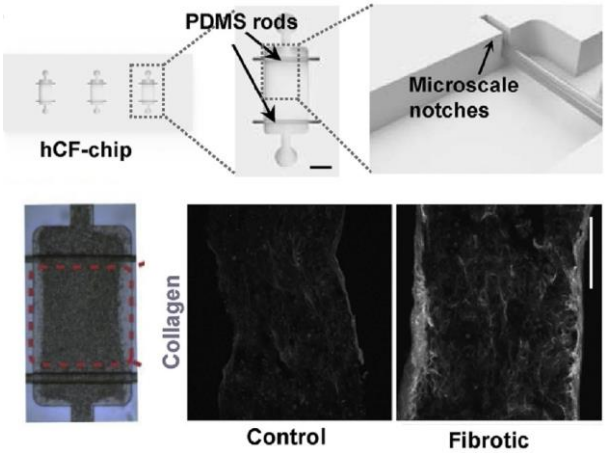
How can we fix this unsightly creation?

- Arrange panels in a way that tells a story
- Spell out acronyms, even if the image in the original article didn't
- Remove the original panel labels and make your own consistent ones
- Label diagrams and plots, even if the image in the original article didn't
- Align the tops and sides of panels; use the same framing style for each panel
- Don't feel compelled to include every bit of detail
- Create your own axis and image labels, and make them consistently legible
- **Have a purpose for including these images rather than simply reproducing them for the sake of showing a picture**

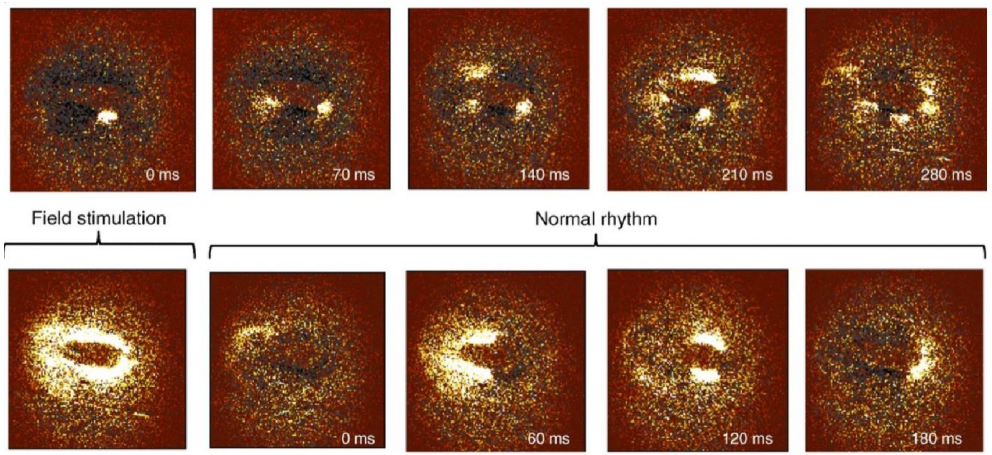
When might it make sense to reproduce images?

- To show what devices or experimental platforms really look like





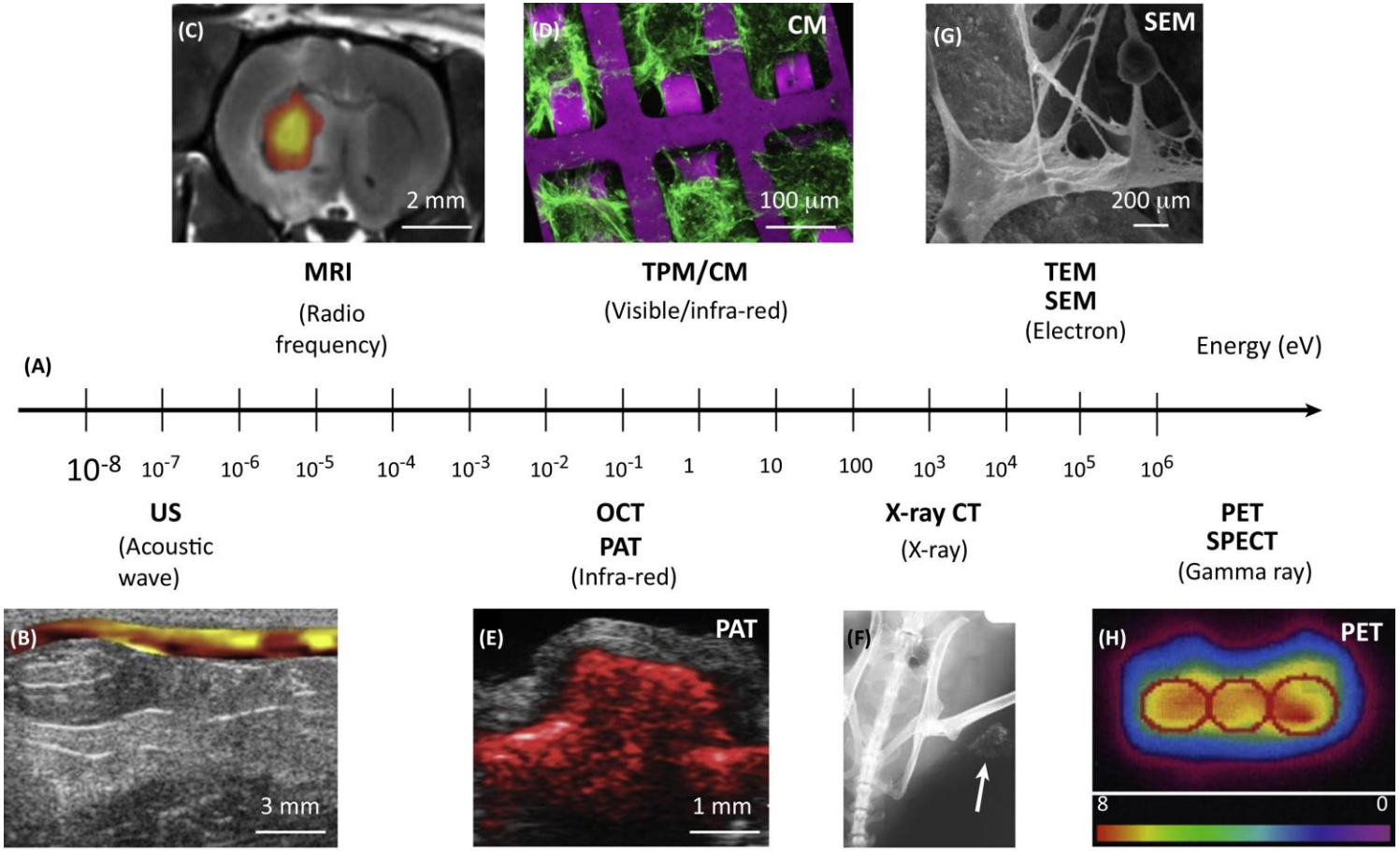
- When the result that you're discussing is itself an image



https://doi.org/10.1016/j.tibtech.2020.08.005

When might it make sense to reproduce images?

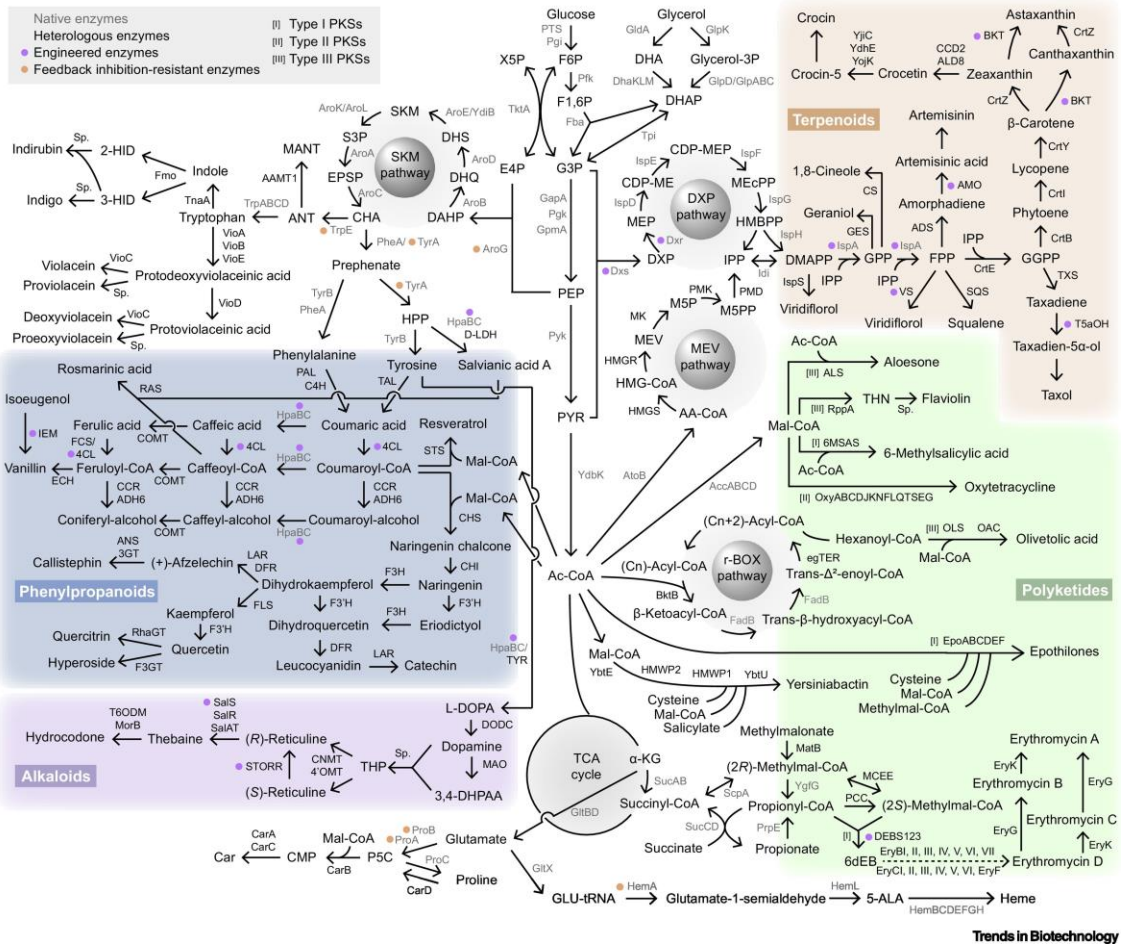
- Better yet, if you can incorporate images from many publications in a way that gives the readers new insight



https://doi.org/10.1016/j.tibtech.2017.09.004

Data collection and replotting

- Combine ideas or results from many publications



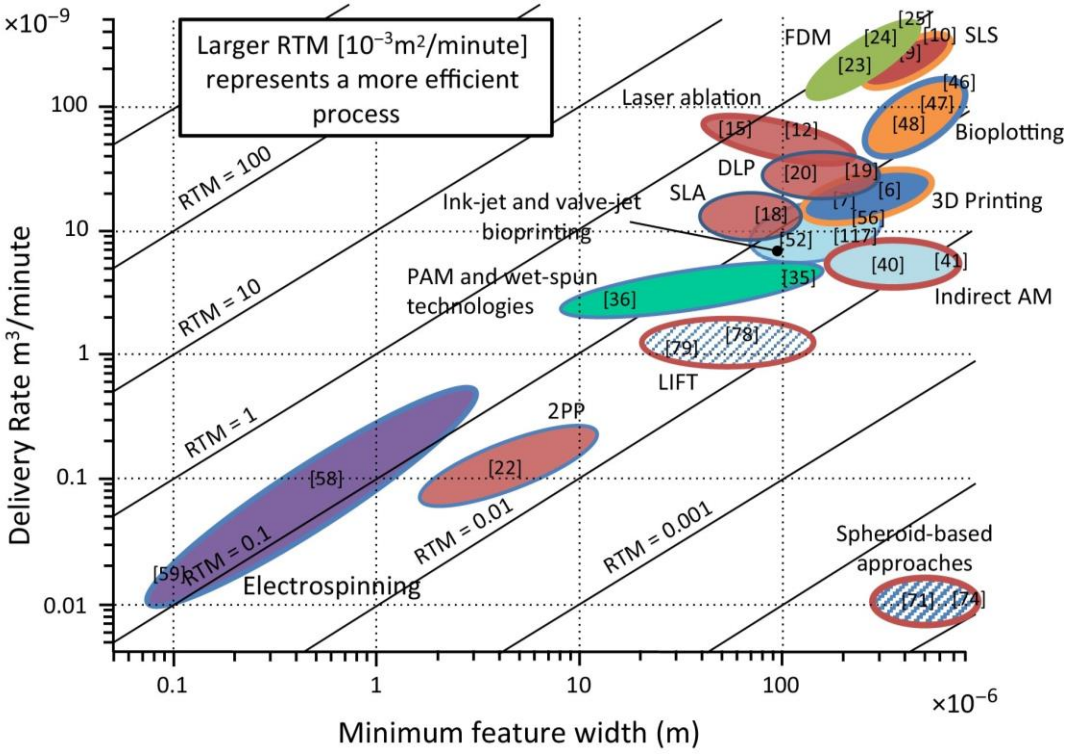
Design language:
Enzymes are color-coded according to their source organism

Pathways that produce compounds with related chemical structure are shown in the same color

https://doi.org/10.1016/j.tibtech.2019.11.007

Data collection and replotting

- Combine ideas or results from many publications, sometimes with new analysis



Design language:
Feature width and delivery rate are commonly reported in this field; this figure plots them against each other to invent a new figure of merit

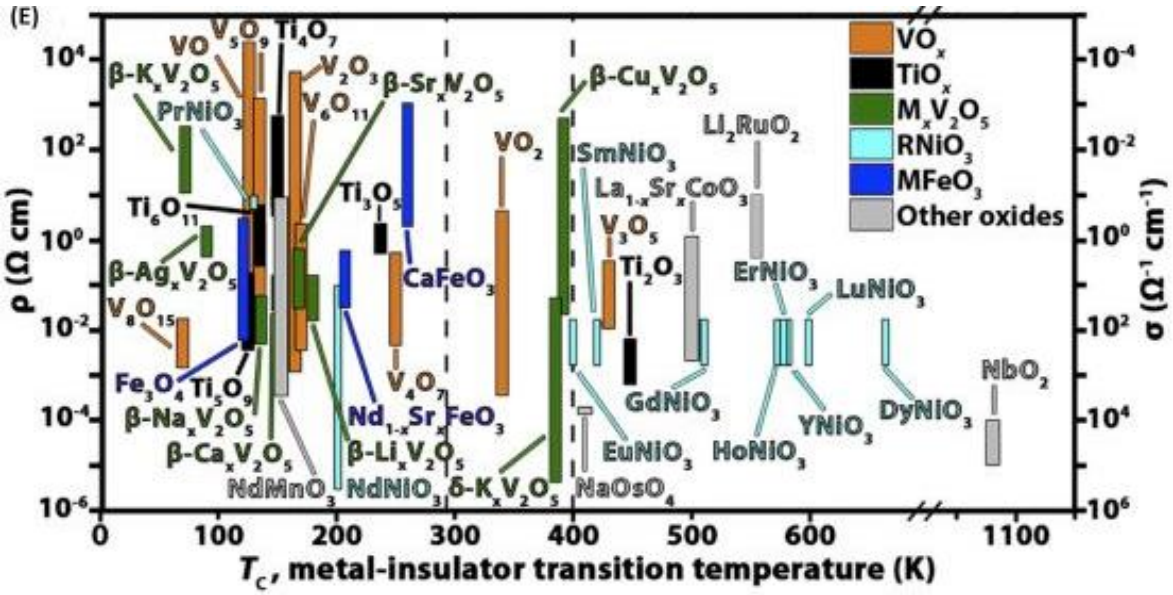
Each number represents a result from one publication

Border color, fill color, and fill style indicate different ways to classify these techniques

https://doi.org/10.1016/j.tibtech.2017.10.015

Data collection and replotting

- Combine ideas or results from many publications, sometimes with new analysis



Design language:
Each bar represents one chemical composition and one or more research results

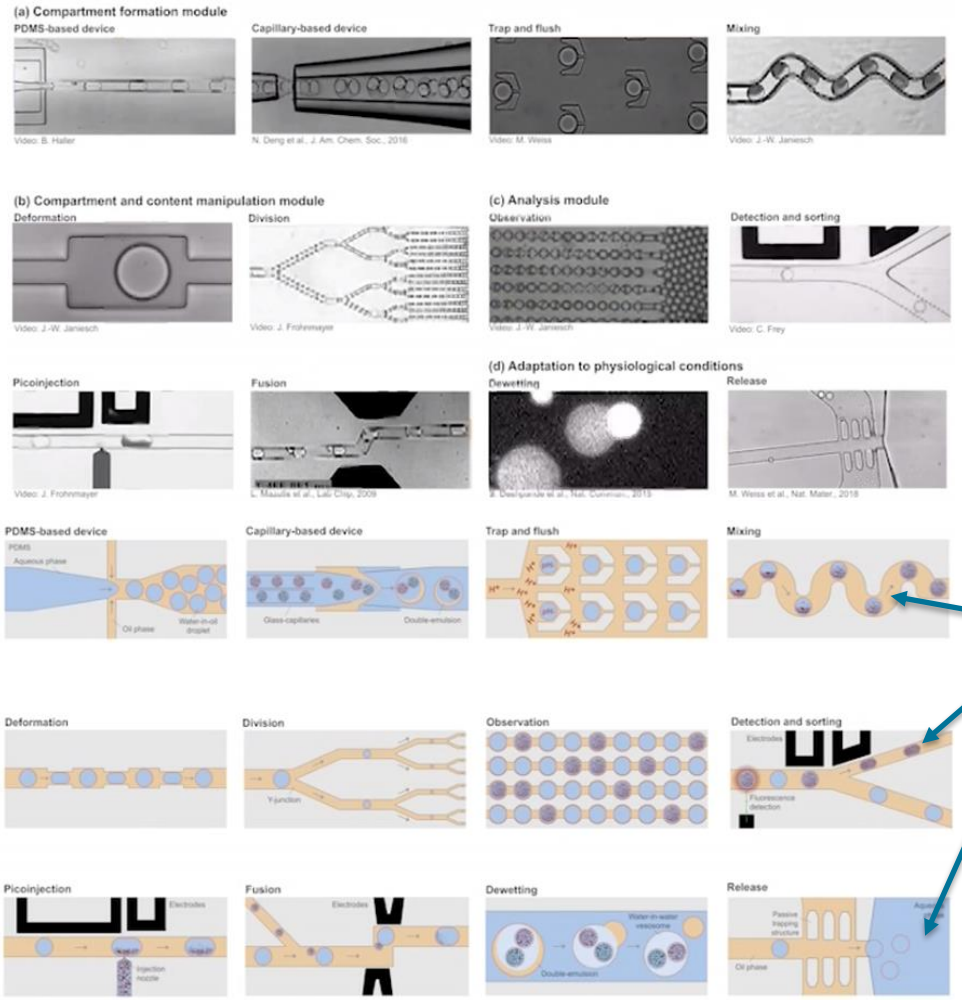
Colors represent distinct classes of materials

Vertical dashed lines represent important thresholds for applications

https://doi.org/10.1016/j.trechm.2019.07.005

Schematic illustrations

- A “cartoon” or diagram that preserves the main idea of a real image

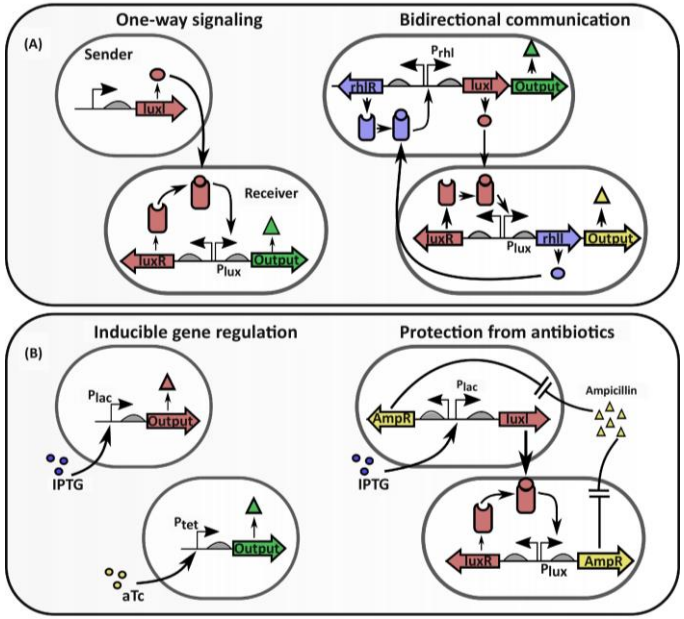


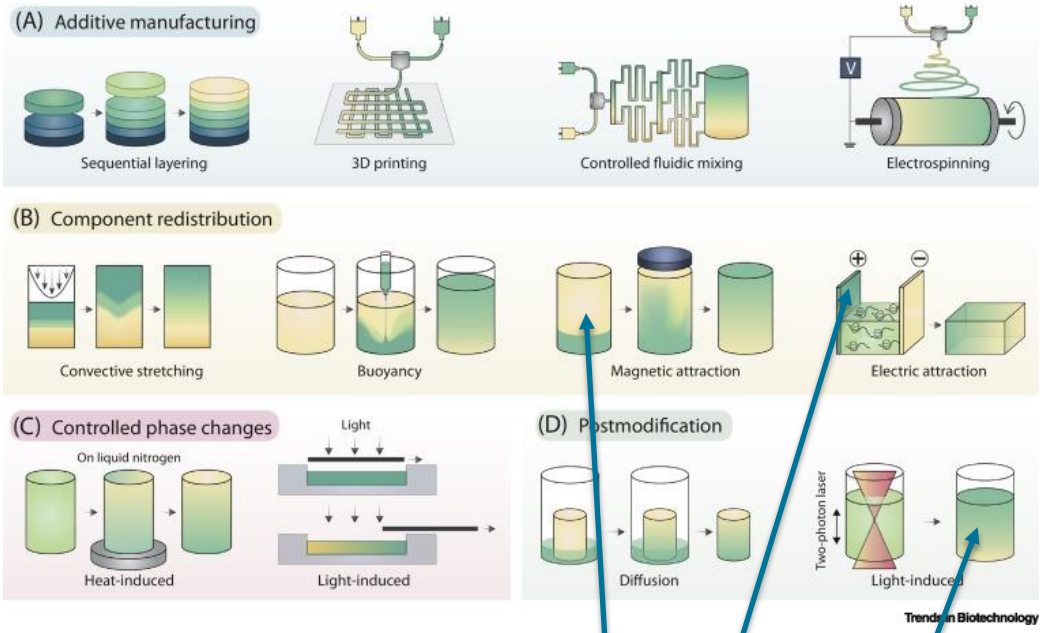
Design language:
Oil phase shown in beige
Biological material shown in purple
Aqueous phase shown in blue

https://doi.org/10.1016/j.tibtech.2018.03.008

Schematic illustrations

- Illustrate a complex concept that is difficult to show a real image of





Design language:
Products, genetic elements, and signaling components from the same pathway are shown in the same color

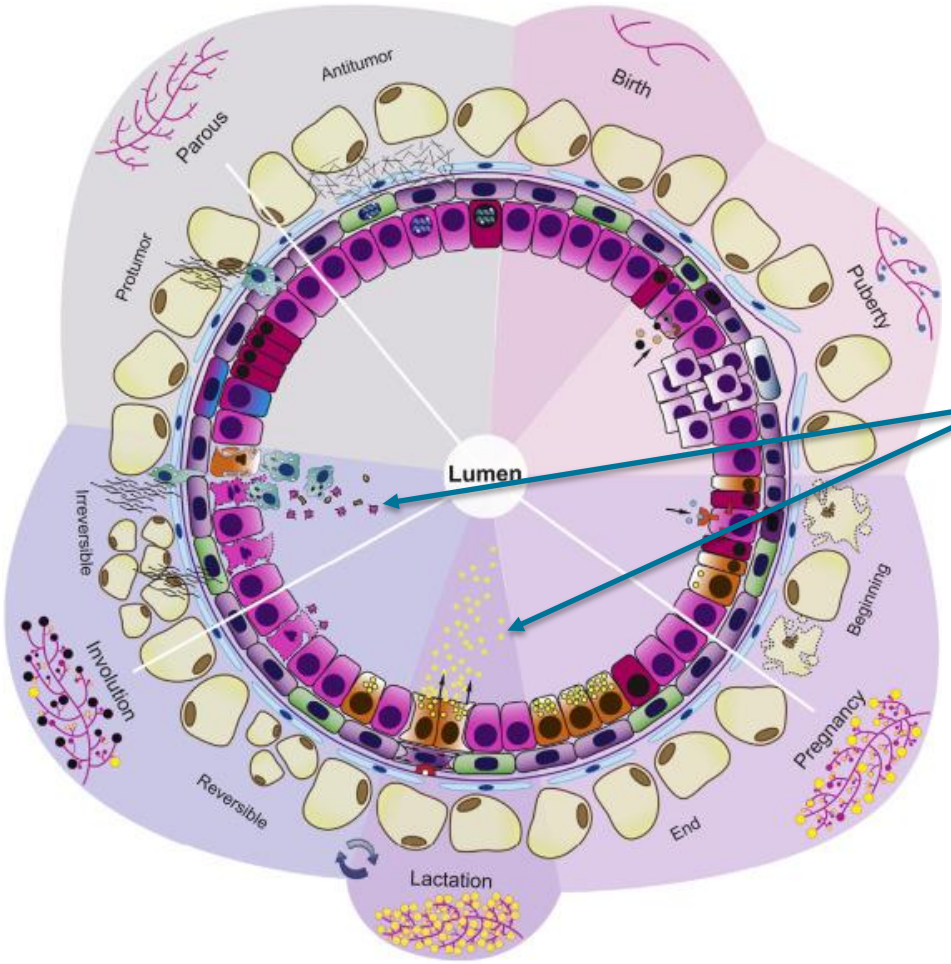
Design language:
Component 1 is shown in yellow
Component 2 is shown in green
Each technique ends with the same image

https://doi.org/10.1016/j.tibtech.2018.11.002

https://doi.org/10.1016/j.tibtech.2020.06.005

Schematic illustrations

- Illustrate a complex concept that is difficult to show a real image of



Design language:
Background color indicates developmental stage

Cell types are the same shape and color across all developmental stages

Graphical elements in “unexpected” locations draw attention to that part of the process

https://doi.org/10.1016/j.molmed.2019.06.003

What do these good figures have in common?

- Add insight beyond just reporting results from other articles, **much like a review article itself should**
- Judiciously reproduce images from prior publications when the image itself is part of the research finding
- Never simply reproduce data like a bar graph, an absorption spectrum, a Kaplan-Meier curve, and so on
- Use consistent colors and symbols throughout the figure—and across multiple figures in the same article
- Take inspiration from real devices, organisms, chemical reactions, etc. to create a simplified visualization of the most important ideas
- Avoid saturated colors and unusual fonts
- Use text to label the images, not tell the story through text

Thank you.

Ask your questions:

Researcher Academy Mendeley group

Follow us on Twitter

See our Cell Mentor blog posts at http://crosstalk.cell.com/blog/frankenfigures-alive-but-easily-misunderstood and http://crosstalk.cell.com/blog/from-mad-scientist-to-artist