How to write a good (math Biology) paper

Eric N. Cytrynbaum
University of British Columbia

Guiding framework

Guiding framework

- Who will be interested in the results? What journal is appropriate? This sets the basic context for writing.

Guiding framework

- Who will be interested in the results? What journal is appropriate? This sets the basic context for writing.
- Scan recent papers in that journal(s) to get a sense of style, format. Journals generally have a distinct culture.

Guiding framework

- Who will be interested in the results? What journal is appropriate? This sets the basic context for writing.
- Scan recent papers in that journal(s) to get a sense of style, format. Journals generally have a distinct culture.
- Here, I'll assume you're writing for a broad (non-math, biological) audience.

Some shoulds and should nots

Some shoulds and should nots

A good paper should

Some shoulds and should nots

A good paper should

- read like a newspaper article, not a mystery novel,

Some shoulds and should nots

A good paper should

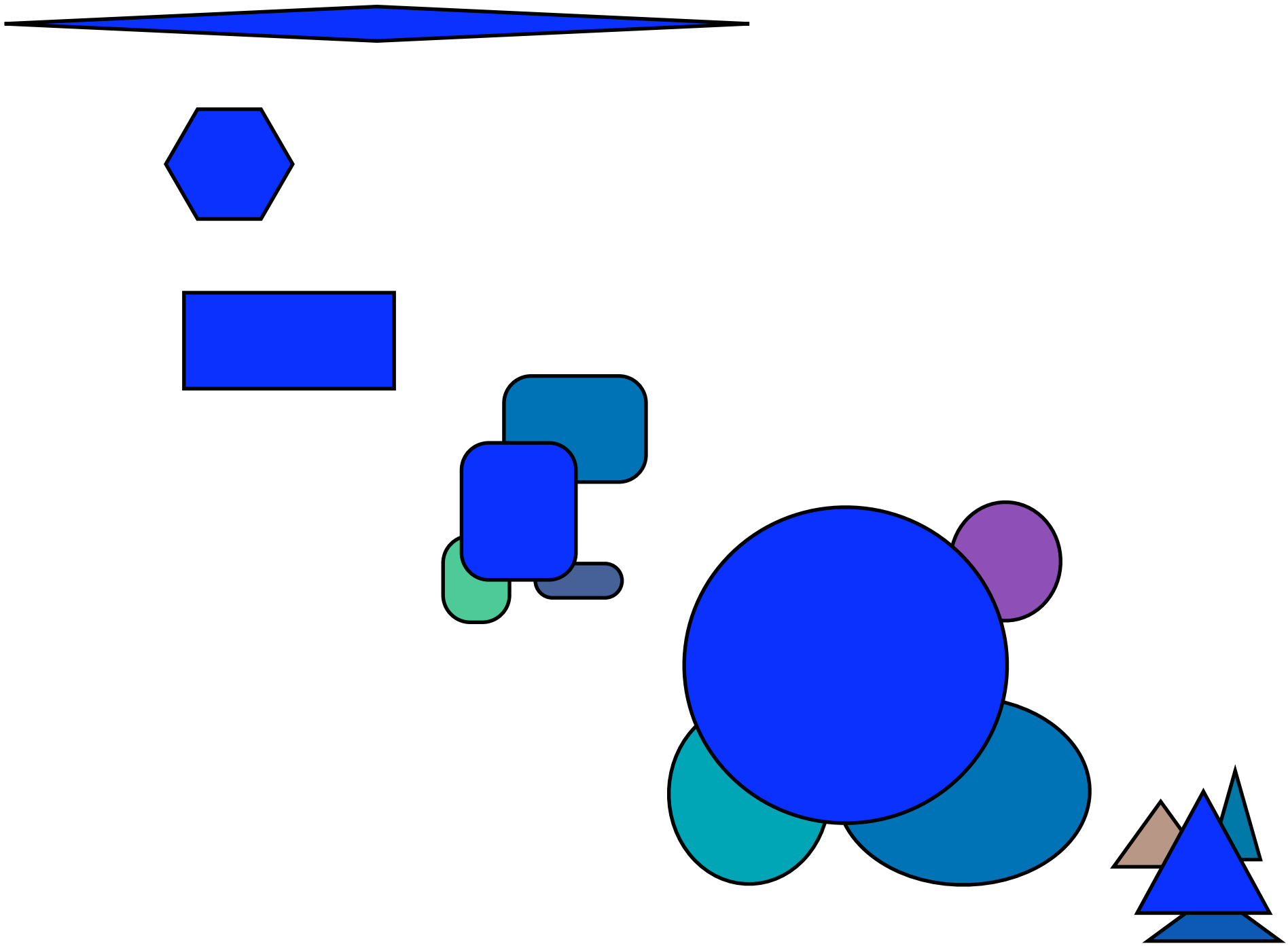
- read like a newspaper article, not a mystery novel,
- be a guided journey through an idea, not a laundry list of loosely related thoughts,

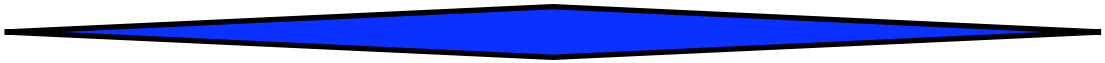
Some shoulds and should nots

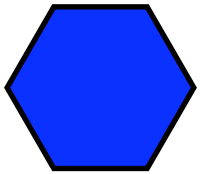
A good paper should

- read like a newspaper article, not a mystery novel,
- be a guided journey through an idea, not a laundry list of loosely related thoughts,
- tell the story the way you wish you had discovered it, not the way you actually did.

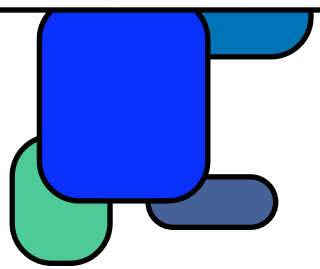
Building the paper

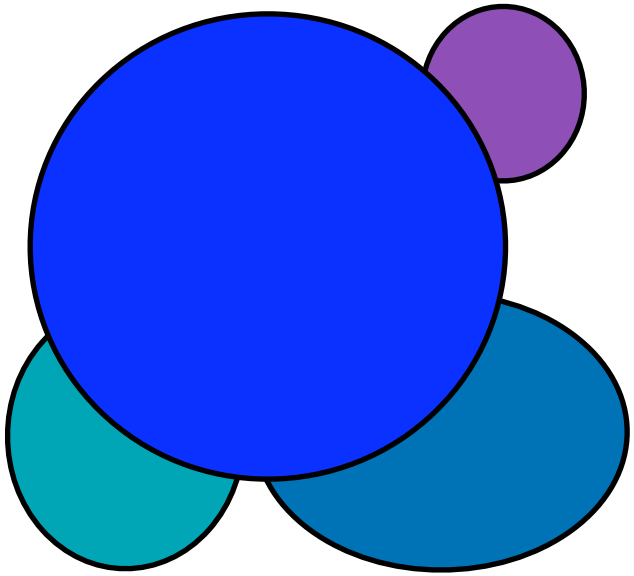


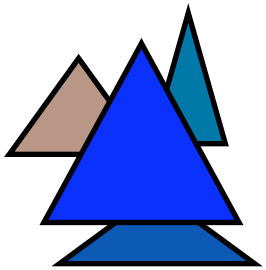


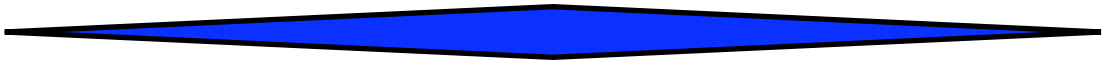


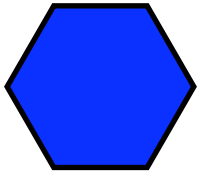
"Bottom line (single message) should be repeated (abstract, intro results, discussion). Everything else, say it once." -- A. Mogilner



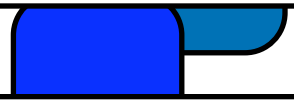




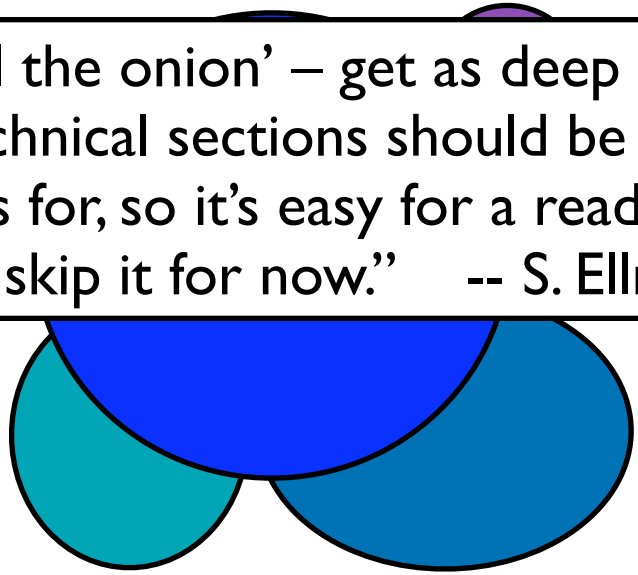


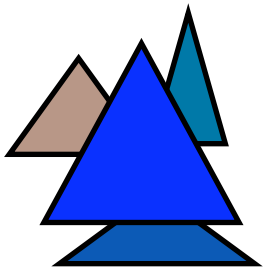


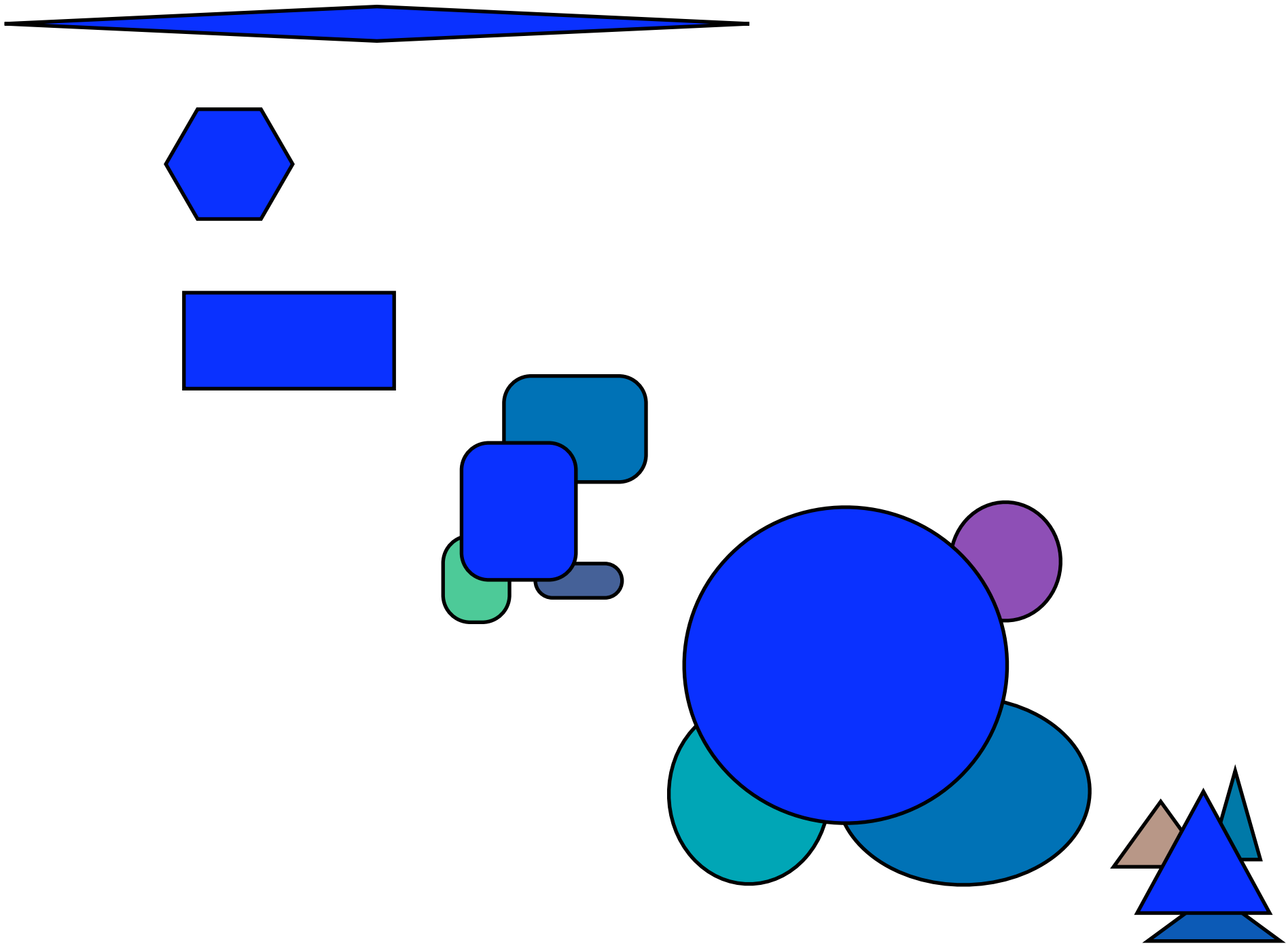
"Bottom line (single message) should be repeated (abstract, intro results, discussion). Everything else, say it once." -- A. Mogilner

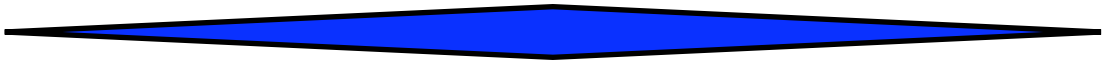


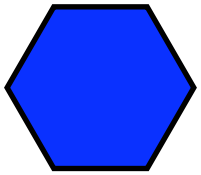
"Use signposting to help people 'peel the onion' – get as deep into the paper as they want, but no deeper. Technical sections should be prefaced by an explanation of what and who it's for, so it's easy for a reader to tell if they should read it, skim it, or skip it for now." -- S. Ellner

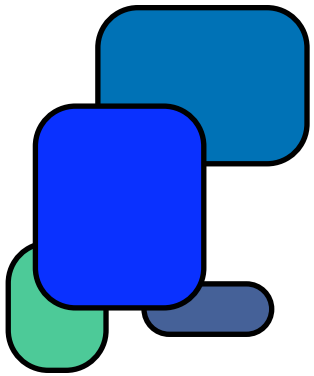




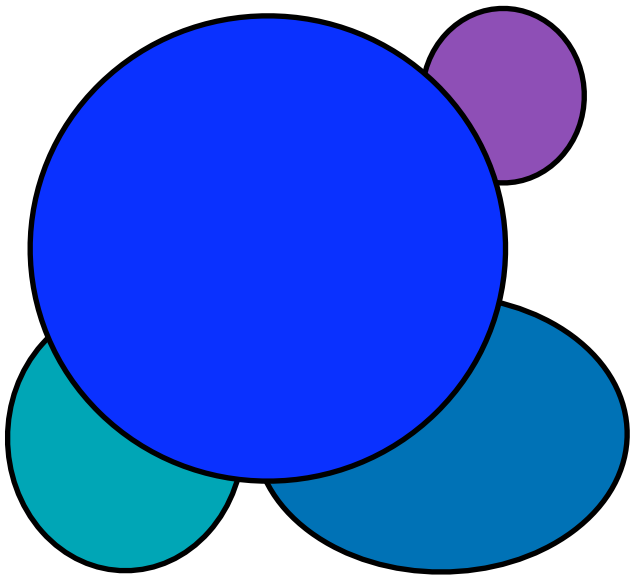


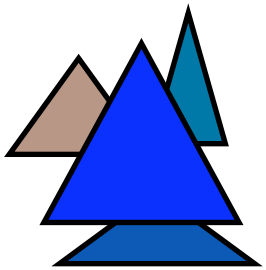


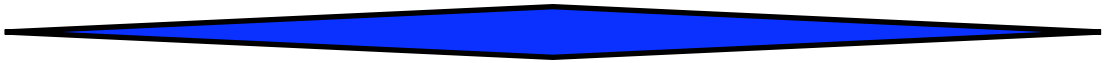


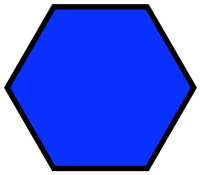


Results (1)

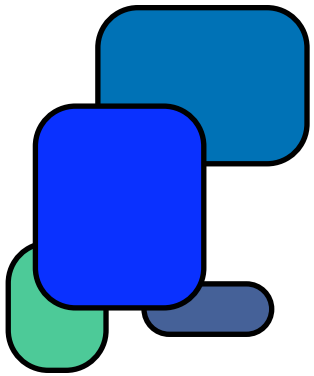




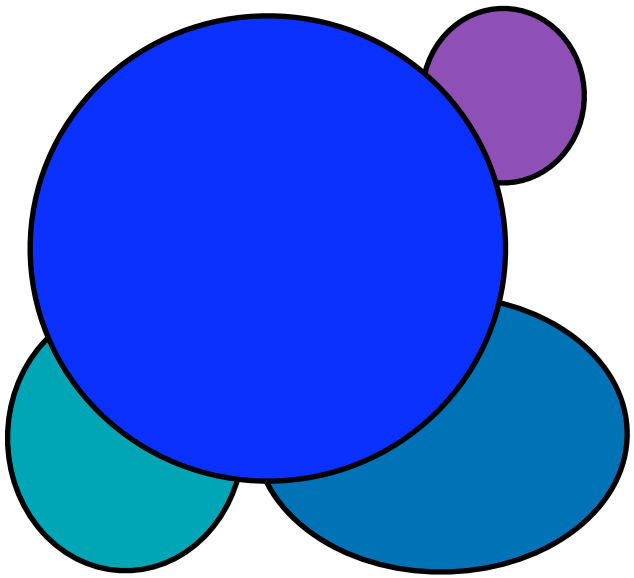


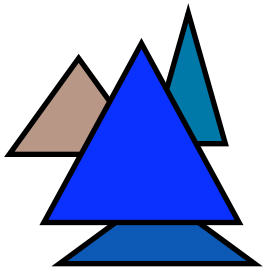


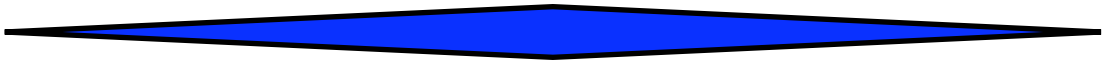
Model (2)

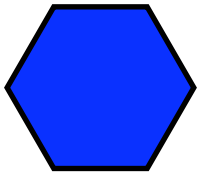


Results (1)

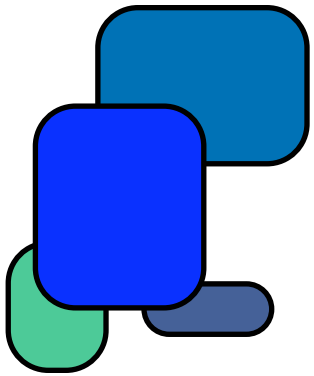




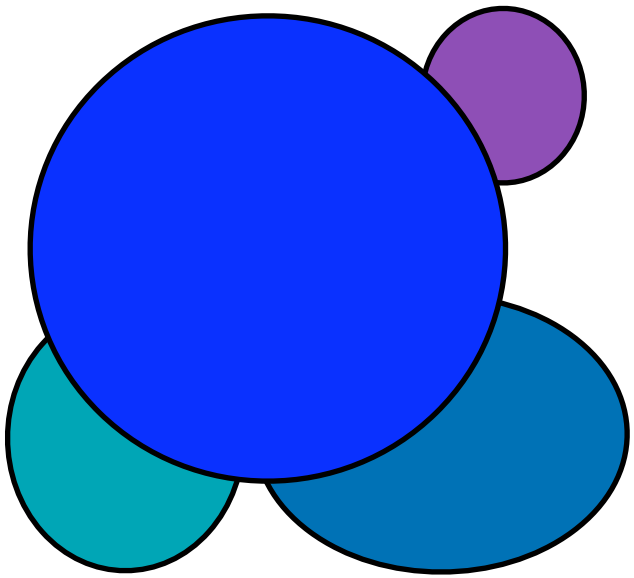


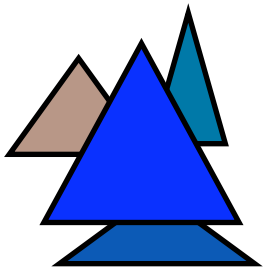


Model (2)

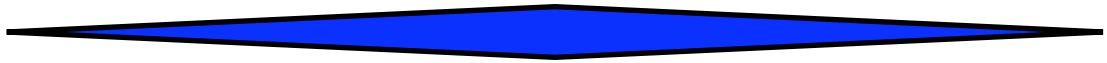


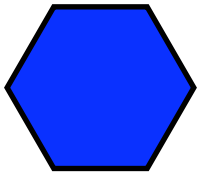
Results (1)





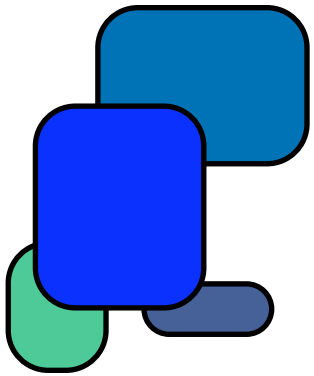
Discussion (3)



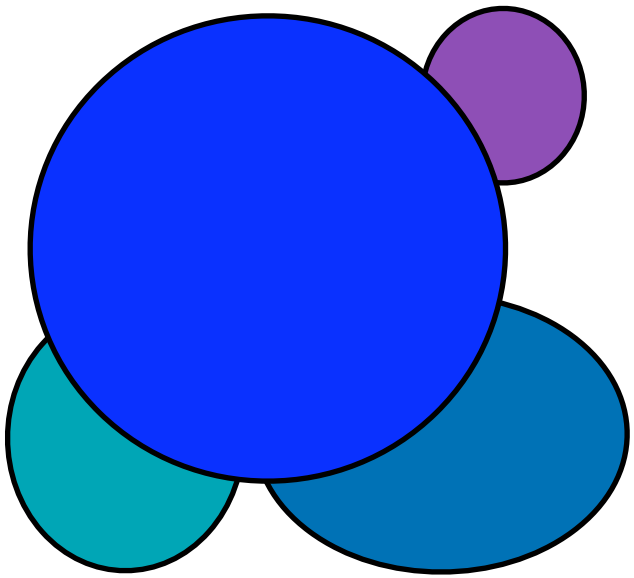


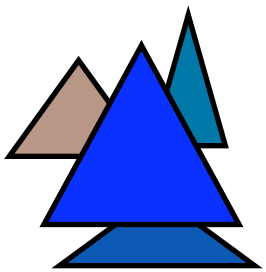
Intro (4)

Model (2)

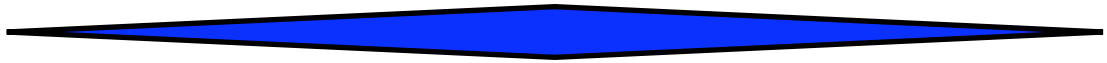


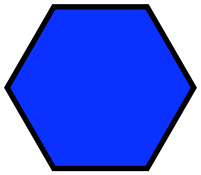
Results (1)





Discussion (3)

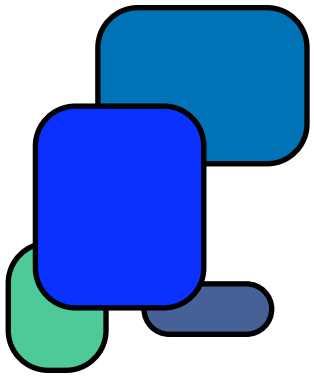




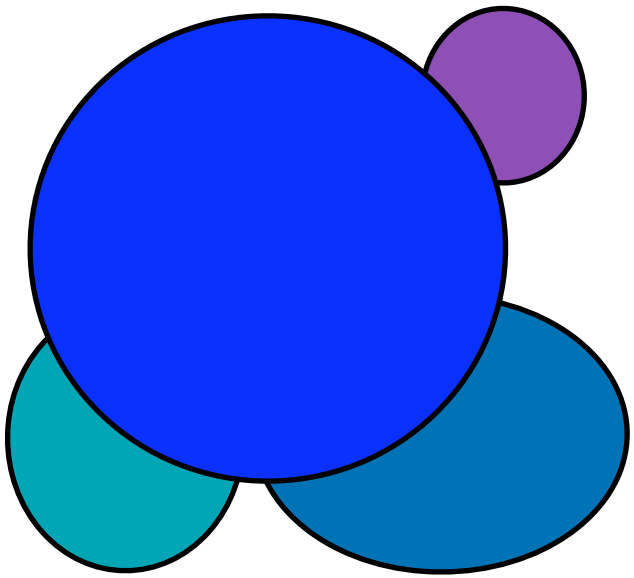
Abstract (5)

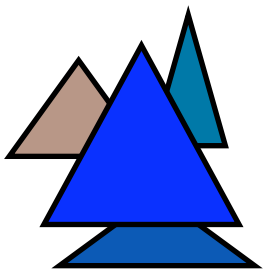
Intro (4)

Model (2)

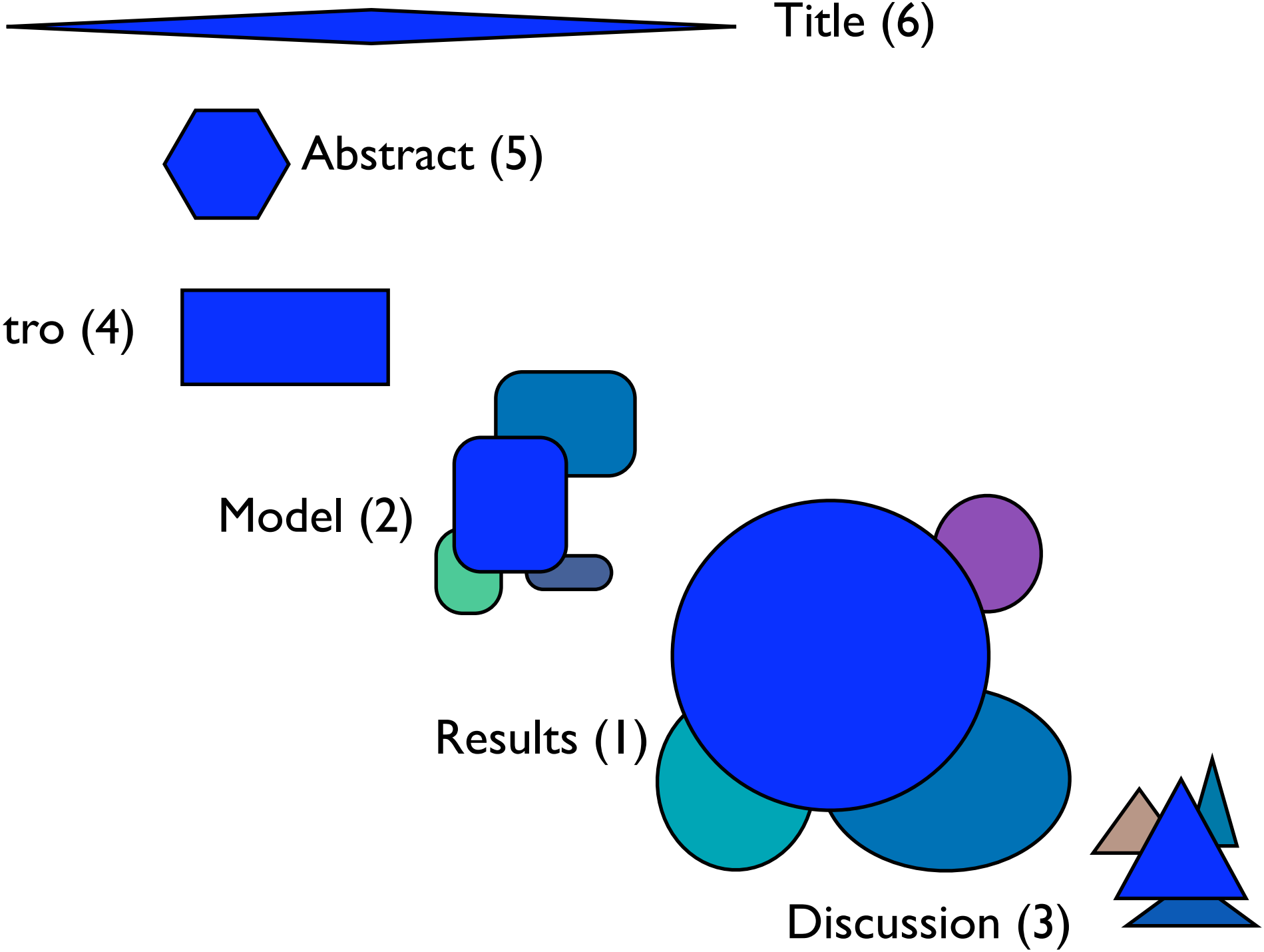


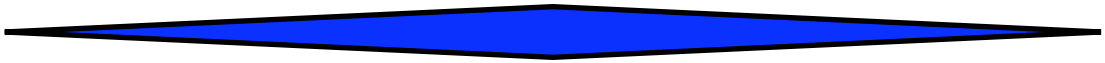
Results (1)



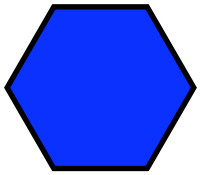


Discussion (3)


Title (6)
Abstract (5)
Intro (4)
Model (2)
Results (1)
Discussion (3)



Title (6)

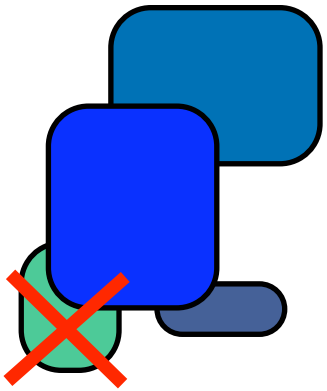


Abstract (5)

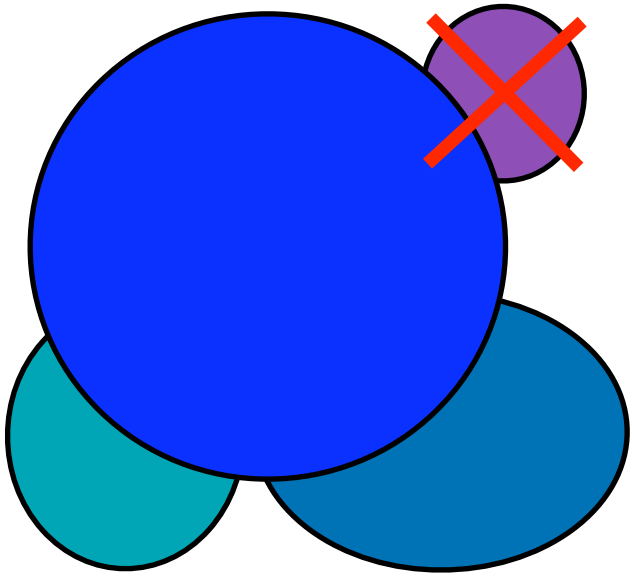
"cut, cut, cut" -- A. Mogilner

Intro (4)

Model (2)



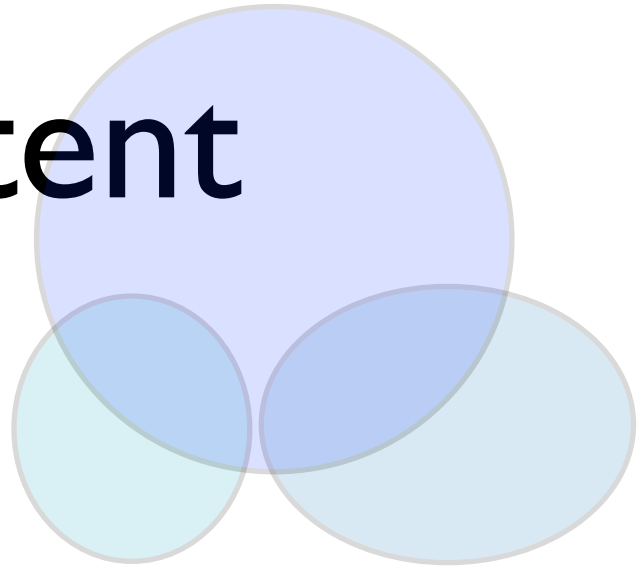
Results (1)



Discussion (3)

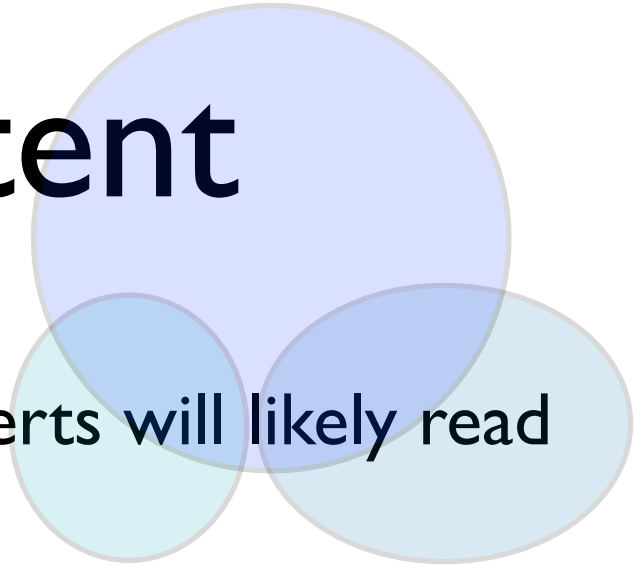
The content

The content



- Results

The content



- Results
  - The core of the paper. Experts will likely read this section only.

The content

- Results
  - The core of the paper. Experts will likely read this section only.
  - Figures and captions should tell much of the story. Lay these out first. Build flow of ideas around these. “Chronological” or most-to-least important.

The content

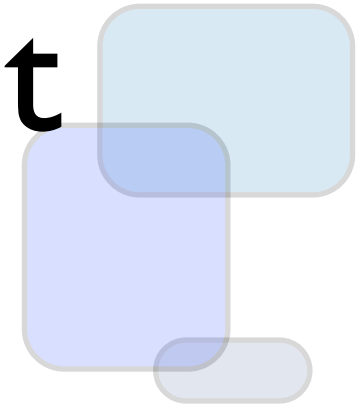
- Results
  - The core of the paper. Experts will likely read this section only.
  - Figures and captions should tell much of the story. Lay these out first. Build flow of ideas around these. “Chronological” or most-to-least important.
  - Must be readable to a non-mathematician. If your results are biologically relevant, they should be communicable without heavy math.

The content

- Results
  - The core of the paper. Experts will likely read this section only.
  - Figures and captions should tell much of the story. Lay these out first. Build flow of ideas around these. "Chronological" or most-to-least important.
  - Must be readable to a non-mathematician. If your results are biologically relevant, they should be communicable without heavy math.
  - "Results" are unambiguous; no interpretation here except for logical transitions between sections.

The content

The content



- Model

The content

- Model
  - Describe in words first, stating assumptions clearly, then equations but only the necessary ones.

The content

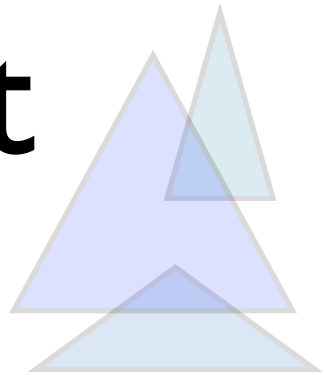
- Model
  - Describe in words first, stating assumptions clearly, then equations but only the necessary ones.
  - Provide enough detail that results can be properly understood and reproduced.

The content

- Model
  - Describe in words first, stating assumptions clearly, then equations but only the necessary ones.
  - Provide enough detail that results can be properly understood and reproduced.
  - Don't cloud the issue with variants irrelevant to Results. Asides interesting to you are distracting to the reader.

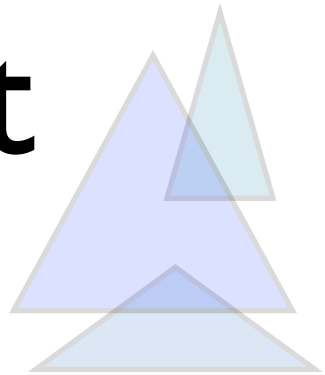
The content

The content



- Discussion

The content



- Discussion
  - Short summary of results (answers to questions and your supporting evidence).

The content

- Discussion
  - Short summary of results (answers to questions and your supporting evidence).
  - Interpretation goes here - tell them what you think the results mean.

The content

- Discussion
  - Short summary of results (answers to questions and your supporting evidence).
  - Interpretation goes here - tell them what you think the results mean.
  - Discuss impact of results (don't repeat results), point out remaining mysteries, highlight non-trivial predictions and broad impact.

The content

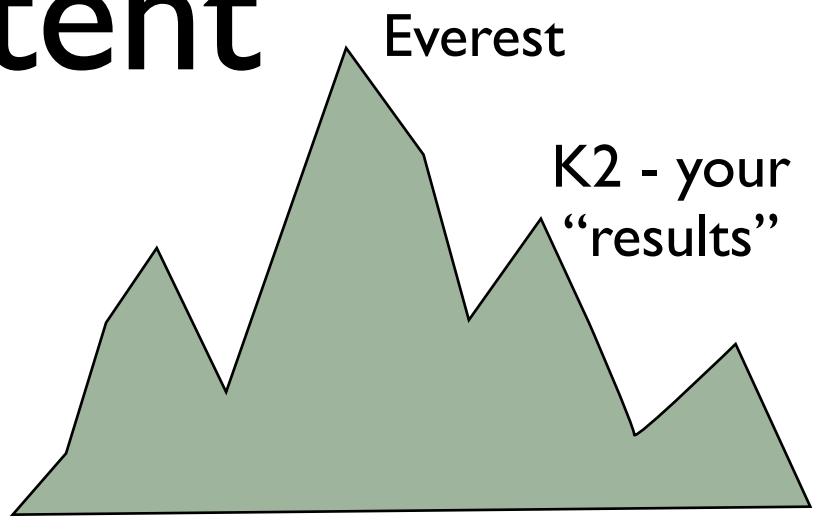
The content

- Introduction

The content

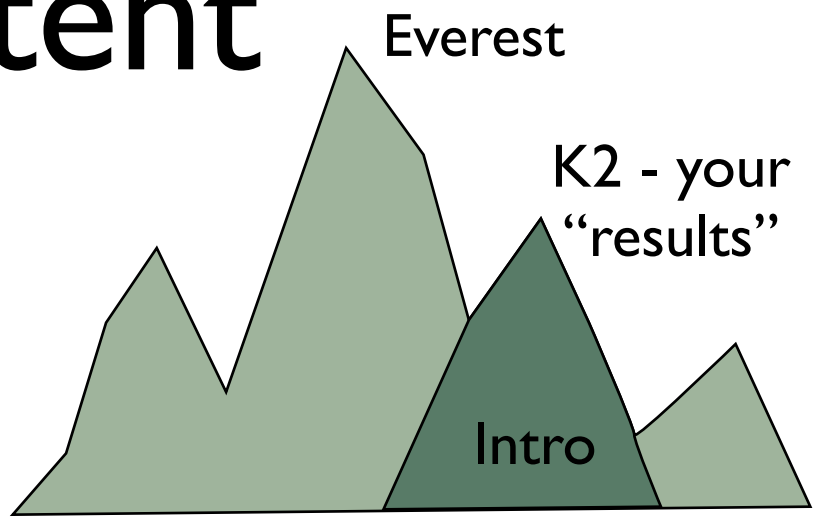
- Introduction
  - Review the relevant history (cite!!) building up to the big remaining unknowns (the topic of your work).

The content



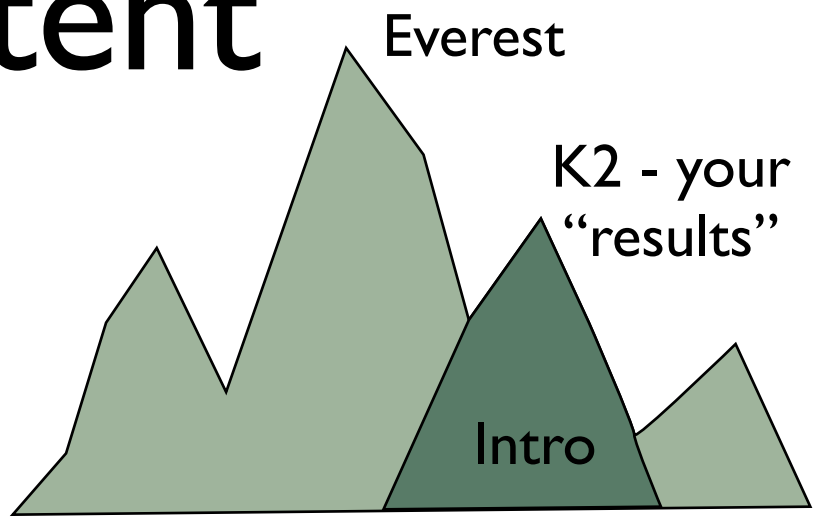
- Introduction
  - Review the relevant history (cite!!) building up to the big remaining unknowns (the topic of your work).

The content



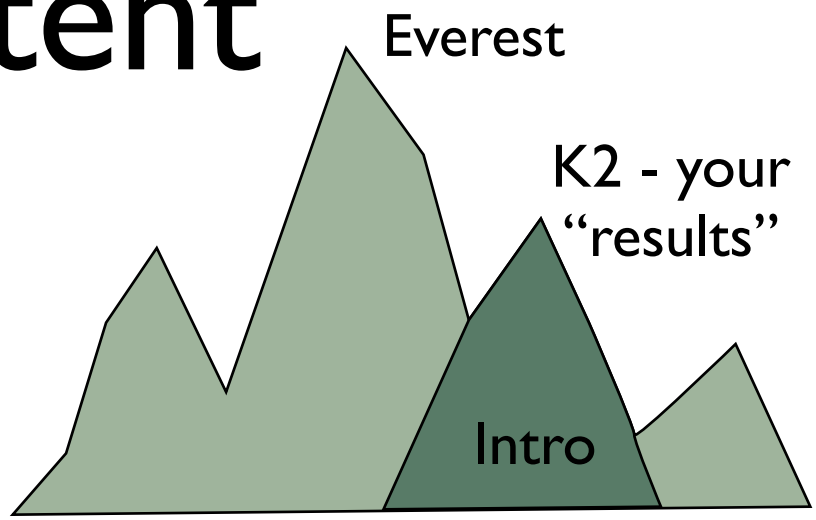
- Introduction
  - Review the relevant history (cite!!) building up to the big remaining unknowns (the topic of your work).

The content



- Introduction
  - Review the relevant history (cite!!) building up to the big remaining unknowns (the topic of your work).
  - State what you've done - leave no mysteries. This section should make the structure of the paper transparent but not explicit (In section 2...).

The content



- Introduction
  - Review the relevant history (cite!!) building up to the big remaining unknowns (the topic of your work).
  - State what you've done - leave no mysteries. This section should make the structure of the paper transparent but not explicit (~~In section 2...~~).

But what about the cool math?

- To a biology audience, the guts of the analysis are for the reviewers.
- Equations necessary for the flow of ideas can be included.
- All else should be appendisized (if necessary to support Results), supplementalized (if helpful in understanding Results) or published elsewhere.

The marketing

The marketing

- Title

The marketing

- Title
  - be specific,

The marketing

- Title
  - be specific,
  - use easily parsed keywords,

The marketing

- Title
  - be specific,
  - use easily parsed keywords,
  - as “big picture” as possible without overstating.

The marketing

- Title
  - be specific,
  - use easily parsed keywords,
  - as “big picture” as possible without overstating.

On the Origin of Species

The marketing

- Title
  - be specific,
  - use easily parsed keywords,
  - as "big picture" as possible without overstating.

~~*On the Origin of Species*~~

Selective survival of fitter genetic variants leads to gradual changes in populations

The marketing

- Title
  - be specific,
  - use easily parsed keywords,
  - as “big picture” as possible without overstating.
- Abstract

The marketing

- Title
  - be specific,
  - use easily parsed keywords,
  - as “big picture” as possible without overstating.
- Abstract
  - what’s the question (1-2 sentences),

The marketing

- Title
  - be specific,
  - use easily parsed keywords,
  - as “big picture” as possible without overstating.
- Abstract
  - what’s the question (1-2 sentences),
  - what’s known already (1-2),

The marketing

- Title
  - be specific,
  - use easily parsed keywords,
  - as “big picture” as possible without overstating.
- Abstract
  - what’s the question (1-2 sentences),
  - what’s known already (1-2),
  - what did you find (2-4) - single message only!!

The marketing

- Title
  - be specific,
  - use easily parsed keywords,
  - as “big picture” as possible without overstating.
- Abstract
  - what’s the question (1-2 sentences),
  - what’s known already (1-2),
  - what did you find (2-4) - single message only!!
  - what’s the broad impact (1)?

1: Nature. 2000 Jan 20;403(6767):335-8.

A synthetic oscillatory network of transcriptional regulators.

Elowitz MB, **Leibler S**.

Department of Molecular Biology and Physics, Princeton University, New Jersey 08544, USA.
melowitz@princeton.edu

Networks of interacting biomolecules carry out many essential functions in living cells, but the 'design principles' underlying the functioning of such intracellular networks remain poorly understood, despite intensive efforts including quantitative analysis of relatively simple systems. Here we present a complementary approach to this problem: the design and construction of a synthetic network to implement a particular function. We used three transcriptional repressor systems that are not part of any natural biological clock to build an oscillating network, termed the repressilator, in Escherichia coli. The network periodically induces the synthesis of green fluorescent protein as a readout of its state in individual cells. The resulting oscillations, with typical periods of hours, are slower than the cell-division cycle, so the state of the oscillator has to be transmitted from generation to generation. This artificial clock displays noisy behaviour, possibly because of stochastic fluctuations of its components. Such 'rational network design may lead both to the engineering of new cellular behaviours and to an improved understanding of naturally occurring networks.

PMID: 10659856 [PubMed - indexed for MEDLINE]

1: Nature. 2000 Jan 20;403(6767):335-8.

A synthetic oscillatory network of transcriptional regulators.

Elowitz MB, Leibler S.

Department of Molecular Biology and Physics, Princeton University, New Jersey 08544, USA.
melowitz@princeton.edu

Networks of interacting biomolecules carry out many essential functions in living cells, but the 'design principles' underlying the functioning of such intracellular networks remain poorly understood, despite intensive efforts including quantitative analysis of relatively simple systems. Here we present a complementary approach to this problem: the design and construction of a synthetic network to implement a particular function. We used three transcriptional repressor systems that are not part of any natural biological clock to build an oscillating network, termed the repressilator, in Escherichia coli. The network periodically induces the synthesis of green fluorescent protein as a readout of its state in individual cells. The resulting oscillations, with typical periods of hours, are slower than the cell-division cycle, so the state of the oscillator has to be transmitted from generation to generation. This artificial clock displays noisy behaviour, possibly because of stochastic fluctuations of its components. Such 'rational network design may lead both to the engineering of new cellular behaviours and to an improved understanding of naturally occurring networks.

PMID: 10659856 [PubMed - indexed for MEDLINE]

1: Nature. 2000 Jan 20;403(6767):335-8.

A synthetic oscillatory network of transcriptional regulators.

Elowitz MB, **Leibler S**.

Department of Molecular Biology and Physics, Princeton University, New Jersey 08544, USA.
melowitz@princeton.edu

Networks of interacting biomolecules carry out many essential functions in living cells, but the 'design principles' underlying the functioning of such intracellular networks remain poorly understood, despite intensive efforts including quantitative analysis of relatively simple systems. Here we present a complementary approach to this problem: the design and construction of a synthetic network to implement a particular function. We used three transcriptional repressor systems that are not part of any natural biological clock to build an oscillating network, termed the repressilator, in Escherichia coli. The network periodically induces the synthesis of green fluorescent protein as a readout of its state in individual cells. The resulting oscillations, with typical periods of hours, are slower than the cell-division cycle, so the state of the oscillator has to be transmitted from generation to generation. This artificial clock displays noisy behaviour, possibly because of stochastic fluctuations of its components. Such 'rational network design may lead both to the engineering of new cellular behaviours and to an improved understanding of naturally occurring networks.

PMID: 10659856 [PubMed - indexed for MEDLINE]

1: Nature. 2000 Jan 20;403(6767):335-8.

A synthetic oscillatory network of transcriptional regulators.

Elowitz MB, **Leibler S**.

Department of Molecular Biology and Physics, Princeton University, New Jersey 08544, USA.
melowitz@princeton.edu

Networks of interacting biomolecules carry out many essential functions in living cells, but the 'design principles' underlying the functioning of such intracellular networks remain poorly understood, despite intensive efforts including quantitative analysis of relatively simple systems. Here we present a complementary approach to this problem: the design and construction of a synthetic network to implement a particular function. We used three transcriptional repressor systems that are not part of any natural biological clock to build an oscillating network, termed the repressilator, in Escherichia coli. The network periodically induces the synthesis of green fluorescent protein as a readout of its state in individual cells. The resulting oscillations, with typical periods of hours, are slower than the cell-division cycle, so the state of the oscillator has to be transmitted from generation to generation. This artificial clock displays noisy behaviour, possibly because of stochastic fluctuations of its components. Such 'rational network design may lead both to the engineering of new cellular behaviours and to an improved understanding of naturally occurring networks.

PMID: 10659856 [PubMed - indexed for MEDLINE]

1: Nature. 2000 Jan 20;403(6767):335-8.

A synthetic oscillatory network of transcriptional regulators.

Elowitz MB, **Leibler S**.

Department of Molecular Biology and Physics, Princeton University, New Jersey 08544, USA.
melowitz@princeton.edu

Networks of interacting biomolecules carry out many essential functions in living cells, but the 'design principles' underlying the functioning of such intracellular networks remain poorly understood, despite intensive efforts including quantitative analysis of relatively simple systems. Here we present a complementary approach to this problem: the design and construction of a synthetic network to implement a particular function. We used three transcriptional repressor systems that are not part of any natural biological clock to build an oscillating network, termed the repressilator, in Escherichia coli. The network periodically induces the synthesis of green fluorescent protein as a readout of its state in individual cells. The resulting oscillations, with typical periods of hours, are slower than the cell-division cycle, so the state of the oscillator has to be transmitted from generation to generation. This artificial clock displays noisy behaviour, possibly because of stochastic fluctuations of its components. Such 'rational network design may lead both to the engineering of new cellular behaviours and to an improved understanding of naturally occurring networks.

PMID: 10659856 [PubMed - indexed for MEDLINE]

1: Science. 2008 May 9;320(5877):792-4. Epub 2008 Apr 17.

Comment in:
Science. 2008 May 9;320(5877):755-6.

Reconstitution of contractile FtsZ rings in liposomes.

Osawa M, **Anderson DE**, **Erickson HP**.

Department of Cell Biology, Duke University Medical Center, Durham, NC 27710-3709, USA.

FtsZ is a tubulin homolog and the major cytoskeletal protein in bacterial cell division. It assembles into the Z ring, which contains FtsZ and a dozen other division proteins, and constricts to divide the cell. We have constructed a membrane-targeted FtsZ (FtsZ-mts) by splicing an amphipathic helix to its C terminus. When mixed with lipid vesicles, FtsZ-mts was incorporated into the interior of some tubular vesicles. There it formed multiple Z rings that could move laterally in both directions along the length of the liposome and coalesce into brighter Z rings. Brighter Z rings produced visible constrictions in the liposome, suggesting that FtsZ itself can assemble the Z ring and generate a force. No other proteins were needed for assembly and force generation.

PMID: 18420899 [PubMed - indexed for MEDLINE]

1: Science. 2008 May 9;320(5877):792-4. Epub 2008 Apr 17.

Comment in:
Science. 2008 May 9;320(5877):755-6.

Reconstitution of contractile FtsZ rings in liposomes.

Osawa M, Anderson DE, Erickson HP.

Department of Cell Biology, Duke University Medical Center, Durham, NC 27710-3709, USA.

FtsZ is a tubulin homolog and the major cytoskeletal protein in bacterial cell division. It assembles into the Z ring, which contains FtsZ and a dozen other division proteins, and constricts to divide the cell. We have constructed a membrane-targeted FtsZ (FtsZ-mts) by splicing an amphipathic helix to its C terminus. When mixed with lipid vesicles, FtsZ-mts was incorporated into the interior of some tubular vesicles. There it formed multiple Z rings that could move laterally in both directions along the length of the liposome and coalesce into brighter Z rings. Brighter Z rings produced visible constrictions in the liposome, suggesting that FtsZ itself can assemble the Z ring and generate a force. No other proteins were needed for assembly and force generation.

PMID: 18420899 [PubMed - indexed for MEDLINE]

1: Science. 2008 May 9;320(5877):792-4. Epub 2008 Apr 17.

Comment in:
Science. 2008 May 9;320(5877):755-6.

Reconstitution of contractile FtsZ rings in liposomes.

Osawa M, **Anderson DE**, **Erickson HP**.

Department of Cell Biology, Duke University Medical Center, Durham, NC 27710-3709, USA.

FtsZ is a tubulin homolog and the major cytoskeletal protein in bacterial cell division. It assembles into the Z ring, which contains FtsZ and a dozen other division proteins, and constricts to divide the cell. We have constructed a membrane-targeted FtsZ (FtsZ-mts) by splicing an amphipathic helix to its C terminus. When mixed with lipid vesicles, FtsZ-mts was incorporated into the interior of some tubular vesicles. There it formed multiple Z rings that could move laterally in both directions along the length of the liposome and coalesce into brighter Z rings. Brighter Z rings produced visible constrictions in the liposome, suggesting that FtsZ itself can assemble the Z ring and generate a force. No other proteins were needed for assembly and force generation.

PMID: 18420899 [PubMed - indexed for MEDLINE]

1: Science. 2008 May 9;320(5877):792-4. Epub 2008 Apr 17.

Comment in:
Science. 2008 May 9;320(5877):755-6.

Reconstitution of contractile FtsZ rings in liposomes.

Osawa M, **Anderson DE**, **Erickson HP**.

Department of Cell Biology, Duke University Medical Center, Durham, NC 27710-3709, USA.

FtsZ is a tubulin homolog and the major cytoskeletal protein in bacterial cell division. It assembles into the Z ring, which contains FtsZ and a dozen other division proteins, and constricts to divide the cell. We have constructed a membrane-targeted FtsZ (FtsZ-mts) by splicing an amphipathic helix to its C terminus. When mixed with lipid vesicles, FtsZ-mts was incorporated into the interior of some tubular vesicles. There it formed multiple Z rings that could move laterally in both directions along the length of the liposome and coalesce into brighter Z rings. Brighter Z rings produced visible constrictions in the liposome, suggesting that FtsZ itself can assemble the Z ring and generate a force. No other proteins were needed for assembly and force generation.

PMID: 18420899 [PubMed - indexed for MEDLINE]

Try it yourself

Try it yourself

- Pick a well-known idea and pretend you're writing the paper in which it first appeared.

Try it yourself

- Pick a well-known idea and pretend you're writing the paper in which it first appeared.
- Choose a title and write an abstract (<250 words).

Try it yourself

- Pick a well-known idea and pretend you're writing the paper in which it first appeared.
- Choose a title and write an abstract (<250 words).
- e.g. Safe and efficient flow of traffic through the use of red, yellow and green lights.

The point of this exercise - to write well you have to think about who and how your work will be read, not just what you want to say. Read papers to learn about form, not just about the science.

Points raised in discussion (BIRS IGTC summit, Sept 20, 2008)

- Avoiding "I" - too much "I" sounds like a personal diary but using passive voice to avoid "I" is stilted and should be avoided.
- Explicit signposting - some like them, others not. They should not replace careful thinking about presentation.
- Referencing and acknowledgments - ideas are not a zero-sum game. Omitting these can be a serious issue.
- Use simple language.
- Authorship: (1) alphabetical - egalitarian but low information content, typical in pure math, (2) descending order of contribution, typical in applied math, (3) descending order of contribution from top, ascending PI contribution from the bottom up, typical in life science.

References and Acknowledgments

- Alex Mogilner (personal exchanges)
- Simon Peyton Jones (web resource)
- Steve Ellner (web posting)
- Sand-Jensen, K. (2007). *How to write consistently boring scientific literature*. Oikos 116:723-727.

Summary

Title: Hey, you, you're interested in reading this!

Abstract: This is why you really don't want to miss reading this. We found out really cool stuff.

Intro: If (and only if) you say it later, give background and set it up here.

Model: Ideas over equations. Those that don't like reading math should still be able to decipher your assumptions.

Results: Organize "chronologically" or most-to-least-important. No fluff. Just the data.

Discussion: This is what we did. This is how it fits in with everything else. This is what it all means. Fluff acceptable here.

Summary

Title: Hey, you, you're interested in reading this!

Abstract: This is why you really don't want to miss reading this. We found out really cool stuff.

Intro: If (and only if) you say it later, give background and set it up here.

Model: Ideas over equations. Those that don't like reading math should still be able to decipher your assumptions.

Results: Organize "chronologically" or most-to-least-important. No fluff. Just the data.

Discussion: This is what we did. This is how it fits in with everything else. This is what it all means. Fluff acceptable here.

I'm done.