Final pitfalls: Submission
• Slimy Simultaneous Submission
  – often detected when same reviewer for both
  – instant dual rejection, often multi-conference blacklisting
• Resubmit Unchanged
  – respond to previous reviews; often get reviewer overlap, irritated if ignored

Later pitfalls: Tactics
• Stealth Contributions
  – don't leave them implicit, it's your job to tell reader explicitly!
  – consider carefully, often different from original project goals
• I Am Utterly Perfect
  – no you're not; discussion of limitations makes paper stronger!

Idiom pitfalls
• Unjustified Visual Encoding
  – should justify why visual encoding design choices appropriate for problem
  – prerequisite: clear statement of problem and encoding!
• Hammer In Search of Nails
  – should characterize capacities of new technique if proposed in paper
• Color Cacophony
  – avoid blatant disregard for basic color perception issues
    – huge areas of highly saturated color
    – categorical color coding for 15+ category levels
    – red/green without luminance differences
    – encoding 3 separate attributes with RGB
• Rainbows Just Like In The Sky
  – avoid hue for ordered attribs, perceptual nonlinearity along rainbow gradient
  – encoding without perceptual encoding awareness

Later pitfalls: Results
• Unfettered By Time
  – choose level of detail for performance numbers
  – detailed graphs for technique papers, high-level for design & eval papers
• Simplicity Comparison
  – compare appropriately against state-of-the-art algorithms
  – head-to-head hardware is best (re-run benchmarks yourself, all on same machine)
• Tiny Toy Datasets
  – compare against state-of-the-art dataset sizes for technique (small ok for eval)
• But My Friends Liked It
  – asking labmates not convincing if target audience is domain experts
• Unjustified Tasks
  – use ecologically valid user study tasks; convincing abstraction of real-world use

Later pitfalls: Strategy
• What I Did Over My Summer Vacation
  – should justify effort rather than contribution
  – don't be too low level, it's not a manual
• Least Publishable Unit
  – avoid tiny increment beyond (your own) previous work
  – make points: new name for old technique
• Dense As Plutonium
  – don't cram in so much content that can't explain why/what/how
  – fails reproducibility test
• Bad Slice and Dice
  – two papers split up wrong
  – neither is standalone, yet both repeat

Final pitfalls: Style
• Deadly Desail Dump
  – explain how only after what and why: provide high-level framing before low-level detail
• Story-Free Captions
  – optimize for flip-through pictures skimming
  – explicitly walk them through images with discussion
• My Picture Speaks For Itself
  – good-low-level flow is necessary (but not sufficient), native speaker check good if ESL
• Grammar Is Optional
  – Mistakes Were Made
  – don't use passive voice, leaves ambiguity about actor
  – your research contribution or done by others?

Generality
• encoding: visualization specific
• strategy: all research
• tactics: all research
• results: visualization specific
• style: all research, except
  – Story-Free Captions, My Picture Speaks For Itself

Contributions in research papers
• what are your research contributions?
  – what can we do that was not possible before?
  – how can we do something better than before?
• what do we know that was unknown or unclear before?
• determines everything
  – from high-level message to which details worth including
  – often not obvious
  – diverged from original goals, in retrospect
  – state them explicitly and clearly in the introduction
  – don't hope reviewer or reader will fill us in for you
  – don't leave unsaid should be obvious after close reading of previous work
• goal is clarity, not overreassuring (limitations typically later, in discussion section)

Idiom pitfalls
• Unjustified Visual Encoding
  – should justify why visual encoding design choices appropriate for problem
  – prerequisite: clear statement of problem and encoding!
• Hammer In Search of Nails
  – should characterize capacities of new technique if proposed in paper
• Color Cacophony
  – avoid blatant disregard for basic color perception issues
    – huge areas of highly saturated color
    – categorical color coding for 15+ category levels
    – red/green without luminance differences
    – encoding 3 separate attributes with RGB
• Rainbows Just Like In The Sky
  – avoid hue for ordered attribs, perceptual nonlinearity along rainbow gradient
  – encoding without perceptual encoding awareness

Later pitfalls: Results
• Unfettered By Time
  – choose level of detail for performance numbers
  – detailed graphs for technique papers, high-level for design & eval papers
• Simplicity Comparison
  – compare appropriately against state-of-the-art algorithms
  – head-to-head hardware is best (re-run benchmarks yourself, all on same machine)
• Tiny Toy Datasets
  – compare against state-of-the-art dataset sizes for technique (small ok for eval)
• But My Friends Liked It
  – asking labmates not convincing if target audience is domain experts
• Unjustified Tasks
  – use ecologically valid user study tasks; convincing abstraction of real-world use

Later pitfalls: Strategy
• What I Did Over My Summer Vacation
  – should justify effort rather than contribution
  – don't be too low level, it's not a manual
• Least Publishable Unit
  – avoid tiny increment beyond (your own) previous work
  – make points: new name for old technique
• Dense As Plutonium
  – don't cram in so much content that can't explain why/what/how
  – fails reproducibility test
• Bad Slice and Dice
  – two papers split up wrong
  – neither is standalone, yet both repeat

Final pitfalls: Style
• Deadly Desail Dump
  – explain how only after what and why: provide high-level framing before low-level detail
• Story-Free Captions
  – optimize for flip-through pictures skimming
• My Picture Speaks For Itself
  – explicitly walk them through images with discussion
• Grammar Is Optional
  – good-low-level flow is necessary (but not sufficient), native speaker check good if ESL
• Mistakes Were Made
  – don't use passive voice, leaves ambiguity about actor
  – your research contribution or done by others?

Contributions in research papers
• what are your research contributions?
  – what can we do that was not possible before?
  – how can we do something better than before?
• what do we know that was unknown or unclear before?
• determines everything
  – from high-level message to which details worth including
  – often not obvious
  – diverged from original goals, in retrospect
  – state them explicitly and clearly in the introduction
  – don't hope reviewer or reader will fill us in for you
  – don't leave unsaid should be obvious after close reading of previous work
• goal is clarity, not overreassuring (limitations typically later, in discussion section)

Research Process & Pitfalls
• Review reading pitfalls
  – Reviewers Were Idiots
    – rare: insufficient background to judge worth
    – if reviewer didn't get your point, many readers won't
  – your job: rewrite so clearly that nobody can misunderstand
• Reviewers Were Threatened By My Brilliance
  – seldom unaided hero since immensely familiar with area
• I Just Know Person X Wrote This Review
  – sometimes true, sometimes false
  – don't get fixated, try not to take it personally
• It's The Writing Not The Work
  – sometimes true bad writing can doom good work (good writing may save borderline)
  – sometimes true: weak work common: reinvent the wheel worse than previous one

Writing InfoVis Papers
• writing infovies papers: pitfalls to avoid
  – Process and Pitfalls in Writing Information Visualization Research Papers.

Idiom pitfalls
• Unjustified Visual Encoding
  – should justify why visual encoding design choices appropriate for problem
  – prerequisite: clear statement of problem and encoding!
• Hammer In Search of Nails
  – should characterize capacities of new technique if proposed in paper
• Color Cacophony
  – avoid blatant disregard for basic color perception issues
    – huge areas of highly saturated color
    – categorical color coding for 15+ category levels
    – red/green without luminance differences
    – encoding 3 separate attributes with RGB
• Rainbows Just Like In The Sky
  – avoid hue for ordered attribs, perceptual nonlinearity along rainbow gradient
  – encoding without perceptual encoding awareness

Later pitfalls: Results
• Unfettered By Time
  – choose level of detail for performance numbers
  – detailed graphs for technique papers, high-level for design & eval papers
• Simplicity Comparison
  – compare appropriately against state-of-the-art algorithms
  – head-to-head hardware is best (re-run benchmarks yourself, all on same machine)
• Tiny Toy Datasets
  – compare against state-of-the-art dataset sizes for technique (small ok for eval)
• But My Friends Liked It
  – asking labmates not convincing if target audience is domain experts
• Unjustified Tasks
  – use ecologically valid user study tasks; convincing abstraction of real-world use

Later pitfalls: Strategy
• What I Did Over My Summer Vacation
  – should justify effort rather than contribution
  – don't be too low level, it's not a manual
• Least Publishable Unit
  – avoid tiny increment beyond (your own) previous work
  – make points: new name for old technique
• Dense As Plutonium
  – don't cram in so much content that can't explain why/what/how
  – fails reproducibility test
• Bad Slice and Dice
  – two papers split up wrong
  – neither is standalone, yet both repeat

Final pitfalls: Style
• Deadly Desail Dump
  – explain how only after what and why: provide high-level framing before low-level detail
• Story-Free Captions
  – optimize for flip-through pictures skimming
• My Picture Speaks For Itself
  – explicitly walk them through images with discussion
• Grammar Is Optional
  – good-low-level flow is necessary (but not sufficient), native speaker check good if ESL
• Mistakes Were Made
  – don't use passive voice, leaves ambiguity about actor
  – your research contribution or done by others?
Final reports
• PDF use InfoVis templates http://junctionpublishing.org/vgtc/Tasks/camera_tvcg.html
– your choice to use LaTeX/Macros
• mid-level discussion of implementation is required – part of your judgement on how much work you did
• high-level: what tools etc did you use
• medium-level: what pre-existing features did you use/adapt
• low-level: detailed manual of how to use, data structure details
• design justification is required – (unless analysis/survey project)
• technique: different in flavour between design study projects and technique projects
• technical experimentation alone is not enough
• publication-level: writing not required – user studies, extensive computational benchmarks, utility to target audience

Course Endgame

Final Presentations Schedule
• 8:00-8:12 Javen Zhang, Matthew Elson, Noam Miron. 
Student Leadership: Making Change Happen in Makerspace Environments.
• 8:12-8:24 David Huang, Sheng Yang, Nathan Luo, Max Sarnowsky. 
Collaborative Visualizations for Community
• 8:24-8:36 Demetrios Dimou, Luana Ponzio, Michael Walker. 
Visualization for Sociable Environments.
• 8:36-8:45 Michelle Langlois, Mike Jonas, Zachar Kudelski. 
Searching Through Maps and Networks.
• 8:45-9:00 Fatma Kursat, Wei Lu, Sofia Strobel, Zhaoran Yu. 
Learning from Graph Patterns.
• 9:00-9:18 Brandon Braun, Zbigniew Szymura, Daniel Kandula, 
Interactive Visual Exploration of Large Datasets. 
• 9:30-9:45 mascot, Project Tearer Image (Project Tearer Image) 
(Final Presentation Slides, PDF)
– template (may change)
– Intro/Framing: 20%
– Main: 30%
– Limitations/Critique/Lessons: 10%
– Slides: 10%
– Presentation/Video Style: 10%
– Question Handling: 10%
• marking by buckets
– great 100%
– good 90%
– ok 78%
– poor 67%
– zero 0%
• marking by buckets
– great 100%
– good 90%
– ok 78%
– poor 67%
– zero 0%
– intro
– main
– critique lessons
– slide style
– presentation video style
– question handling
• 36% Aspnic Assessment
– 9 week, 4 per week
– 75% own comments, 25% responses
• 2% final presentations
– almost all got full credit
• 12 sessions, 1% per session
– 2% final presentations
– (almost all got full credit)
• Final Presentations, Wed Dec 15 2-5pm
– slides/video upload
– upload to Canvas Assignments: Final Slides (mandatory), Final Video (optional)
– – no additional work on project allowed after presentation deadline
– – additional two days to get it all written down coherently for final report

Course structure
• 50% Project, summative assessment at end
– Intro/Framing: 20%
– Main: 30%
– Limitations/Critique/Lessons: 10%
– Slides: 10%
– Presentation/Video Style: 10%
– Question Handling: 10%
• marking by buckets
– great 100%
– good 90%
– ok 78%
– poor 67%
– zero 0%
• marking by buckets
– great 100%
– good 90%
– ok 78%
– poor 67%
– zero 0%
– intro
– main
– critique lessons
– slide style
– presentation video style
– question handling
• 36% Aspnic Assessment
– 9 week, 4 per week
– 75% own comments, 25% responses
• 2% final presentations
– almost all got full credit
• 12 sessions, 1% per session
– 2% final presentations
– (almost all got full credit)
• Final Presentations, Wed Dec 15 2-5pm
– slides/video upload
– upload to Canvas Assignments: Final Slides (mandatory), Final Video (optional)
– – no additional work on project allowed after presentation deadline
– – additional two days to get it all written down coherently for final report

Sample outlines: Design study
• www.cs.ubc.ca/~tmm/courses/547-21/projectdesc.html

Course requirements vs research paper standards
• research novelty not required
• mid-level discussion of implementation is required – part of your judgement on how much work you did
• high-level: what tools etc did you use
• medium-level: what pre-existing features did you use/adapt
• low-level: detailed manual of how to use, data structure details
• design justification is required – (unless analysis/survey project)
• technique: different in flavour between design study projects and technique projects
• technical experimentation alone is not enough
• publication-level: not required – user studies, extensive computational benchmarks, utility to target audience

Paper writing process suggestions
• pre-paper talk
– write and give talk first, as if presenting at conference
– iterate on talk slides to get structure, ordering, arguments right
– then create paper outline from final draft of slides
• encourages concise explanations of critical ideas, creation of key diagrams
• concise summary of research, argumentation and rhetoric
• easier to cut slides that you've already argued over
• pre-paper practice talk feedback session: at least 2-3x talk length
• global comments, then slide by slide detailed discussion
• nurture culture of internal critique (build your own critique group if necessary)
• have non-authors read paper before submitting
– internal review can catch many problems
• ideally group feedback session as above

Course Endgame

Final Reports

Logistics
• Assignments: Final Presentations on Canvas
– upload due Wed Dec 15 noon (2 hrs before session)
– required but not posted: code incl README (Project Source Code and Other Materials, zip)
– encouraged & posted: showcasing image (Project Tearer Image, png)
– required & posted: slides (Project Final Presentation Slides, PDF)
– required but not posted code incl README (Project Source Code and Other Materials, zip)
– encouraged & posted links to all project pages to browse 2020-2003
– don’t compare against paper you would have written
• review the paper they submitted

Course Endgame

Review writing pitfalls
• Uncalibrated Dissonance
– remember you’ve only read the best of the best!
– most new reviewers are overly harsh
• It’s Been Done, Full Stop
– you must say who did it in which paper, full citation is best
• You Didn’t Cite Me
– stop and think whether it’s appropriate
• You Didn’t Channel Me
– don’t compare against paper you would have written
• review the paper they submitted

Course Endgame

Conference talk pitfalls
• Results As Satirical
– don’t waste until the end as a reward for the stalwarts!
– showcase early to motivate
• A Thousand Words, No Pictures
– aggressively replace words with illustrations
– meat slides should have a picture
• Full Coverage Or Bust
– can’t fit all details from paper
– communicate BIG picture
– talk as advertising: convince them it’s worth their time to read paper

Course Endgame

Final presentations marking
• template (may change)
– Intro/Framing: 20%
– Main: 30%
– Limitations/Critique/Lessons: 10%
– Slides: 10%
– Presentation/Video Style: 10%
– Question Handling: 10%
• marking by buckets
– great 100%
– good 90%
– ok 78%
– poor 67%
– zero 0%
• marking by buckets
– great 100%
– good 90%
– ok 78%
– poor 67%
– zero 0%
– intro
– main
– critique lessons
– slide style
– presentation video style
– question handling
• 36% Aspnic Assessment
– 9 week, 4 per week
– 75% own comments, 25% responses
• 2% final presentations
– almost all got full credit
• 12 sessions, 1% per session
– 2% final presentations
– (almost all got full credit)
Preparing your research for publication:

- **Introduction**: Summarize the background and motivate the importance of your study.
- **Related Work**: Discuss how your work compares to existing research.
- **Methodology**: Detail the methods and tools you used.
- **Results**: Present the findings of your study.
- **Discussion**: Interpret the results and relate them to the existing literature.
- **Conclusion**: Summarize the main findings and their implications.
- **Bibliography**: List all the references used in your research.

**Marking**

- **demonstrate design and technique & explainer**: 12.5% each for
  - intro
  - related work
  - solutions
  - implementation/milestones
  - results
  - discussion

- **10% style, 2.5% bibliography**

**Open Science**

- **Available, Reproducible, & Replicable Research**

- **Code & Video**
  - required submit your code
  - see what you've done, I will not post
  - include README file at root with brief roadmap/overview of organization
  - which parts are your code vs libraries
  - how to compile & run
  - do not necessarily expect your code compiles on my machine

- **encouraged but not required**
  - submit live demo URL (provide in README.txt file)
  - open-source your code (if not, to just send me that URL)
  - submit supporting video (if different from final presentation)
  - with or without voiceover
  - very nice to have, software bistro makes demos not last forever!

**Evaluations**

- **showcase image for projects page**
  - 300x300 image
  - call it showcase.png

- **Marking**
  - design study & technique & explainer
  - 12.5% each for
    - intro
    - related work
    - solutions
    - implementation/milestones
    - results
    - discussion

- **10% style, 2.5% bibliography**

**Disseminating research**

- **paper page for each paper**
- **presentation**
  - make sure it's online
  - make sure it's accessible when you move on to new place
  - external archives are better yet (arXiv)

- **algorithm**
  - well documented in paper itself
  - document further with supplemental materials

- **code**
  - make available as open source
  - pick right spot on continuum of efforts involved, from minimal to massive
  - just put it up online and do minimal documentation
  - well documented and tested
  - build a whole community - not the common case

**Reproducibility: Levels to consider**

- **paper**
  - 0: out of the question
  - 1: just put it online
  - 2: make sure it's accessible when you move on to new place
  - 3: consider attending to arXiv
  - 4: 15 minutes with proprietary tools
  - 5: minimal documentation
  - 6: open source code

- **Open Science**
  - **Available, Reproducible, & Replicable Research**

- **Code & Video**
  - required submit your code
  - see what you've done, I will not post
  - include README file at root with brief roadmap/overview of organization
  - which parts are your code vs libraries
  - how to compile & run
  - do not necessarily expect your code compiles on my machine

- **encouraged but not required**
  - submit live demo URL (provide in README.txt file)
  - open-source your code (if not, to just send me that URL)
  - submit supporting video (if different from final presentation)
  - with or without voiceover
  - very nice to have, software bistro makes demos not last forever!

**Evaluations**

- **showcase image for projects page**
  - 300x300 image
  - call it showcase.png

- **Marking**
  - design study & technique & explainer
  - 12.5% each for
    - intro
    - related work
    - solutions
    - implementation/milestones
    - results
    - discussion

- **10% style, 2.5% bibliography**

**Disseminating research**

- **paper page for each paper**
- **presentation**
  - make sure it's online
  - make sure it's accessible when you move on to new place
  - external archives are better yet (arXiv)

- **algorithm**
  - well documented in paper itself
  - document further with supplemental materials

- **code**
  - make available as open source
  - pick right spot on continuum of efforts involved, from minimal to massive
  - just put it up online and do minimal documentation
  - well documented and tested
  - build a whole community - not the common case

**Reproducibility: Levels to consider**

- **paper**
  - 0: out of the question
  - 1: just put it online
  - 2: make sure it's accessible when you move on to new place
  - 3: consider attending to arXiv
  - 4: 15 minutes with proprietary tools
  - 5: minimal documentation
  - 6: open source code

- **Open Science**
  - **Available, Reproducible, & Replicable Research**

- **Code & Video**
  - required submit your code
  - see what you've done, I will not post
  - include README file at root with brief roadmap/overview of organization
  - which parts are your code vs libraries
  - how to compile & run
  - do not necessarily expect your code compiles on my machine

- **encouraged but not required**
  - submit live demo URL (provide in README.txt file)
  - open-source your code (if not, to just send me that URL)
  - submit supporting video (if different from final presentation)
  - with or without voiceover
  - very nice to have, software bistro makes demos not last forever!

**Evaluations**

- **showcase image for projects page**
  - 300x300 image
  - call it showcase.png

- **Marking**
  - design study & technique & explainer
  - 12.5% each for
    - intro
    - related work
    - solutions
    - implementation/milestones
    - results
    - discussion

- **10% style, 2.5% bibliography**

**Disseminating research**

- **paper page for each paper**
- **presentation**
  - make sure it's online
  - make sure it's accessible when you move on to new place
  - external archives are better yet (arXiv)

- **algorithm**
  - well documented in paper itself
  - document further with supplemental materials

- **code**
  - make available as open source
  - pick right spot on continuum of efforts involved, from minimal to massive
  - just put it up online and do minimal documentation
  - well documented and tested
  - build a whole community - not the common case

**Reproducibility: Levels to consider**

- **paper**
  - 0: out of the question
  - 1: just put it online
  - 2: make sure it's accessible when you move on to new place
  - 3: consider attending to arXiv
  - 4: 15 minutes with proprietary tools
  - 5: minimal documentation
  - 6: open source code

- **Open Science**
  - **Available, Reproducible, & Replicable Research**

- **Code & Video**
  - required submit your code
  - see what you've done, I will not post
  - include README file at root with brief roadmap/overview of organization
  - which parts are your code vs libraries
  - how to compile & run
  - do not necessarily expect your code compiles on my machine

- **encouraged but not required**
  - submit live demo URL (provide in README.txt file)
  - open-source your code (if not, to just send me that URL)
  - submit supporting video (if different from final presentation)
  - with or without voiceover
  - very nice to have, software bistro makes demos not last forever!
Replication: crisis in psychology, medicine, etc
• early rumblings left me with (ignorable) qualms
  – papers: is most published research false. Sorkin Deliver Babies (p = 0.008). The Earth is spherical (p < 0.05). False-Positive Psychology
• groundswell of change for what methods are considered legitimate
  – out: ORFs (questionable research practices)
    • p-hacking / p-value fishing / data dredging
    • Hypothesizing After Results are Known (HARKing)
    – in
    • replication
    • pre-registration
    • breaths with biomodal responses
    • some people doubting down and defending previous work
    • many willing to replicate (their own) earlier styles of work

Remarkable introspection on methods
• thoughtful willingness to change standards of field
  – Andrew Gelman’s commentary on the Susan Fiske article
  – Simine Vazire’s entire Sometimes I’m Wrong blog
  – http://sometimesimwrong.typepad.com/
  – especially posts on topic Scientific Integrity
  – Joe Simmons Data Colada blog post What I Want Our Field to Prioritize
  – http://datacolada.org/31/
  – Dana Carvey’s brain statement on her previous power pose work

Visual Design Process In Depth: Dear Data
• inspiring celebration of data humanism
• Giorgia Lupi and Stefanie Posavec

http://www.dear-data.com/by-week/
http://www.datasketch.es/
http://www.makeovermonday.co.uk/blog/

Reproducibility: Levels to consider, cont.
• data
  – make available
  – technical/algorithm data used by system
  – openly share visualization data might not be yours to release
• evaluation: user study results
  – online approval / IRB (personally identifiable information) sanitized; needs advance planning
• parameters
  – how exactly to regenerate/reproduce figures, tables
  – example: http://www.cs.utah.edu/~gk/papers/vis03/

When and how will this storm hit visualization?
• they’re ahead of us
  – they have some paper retractions
  – we don’t (yet) have any retractions for methodological considerations
  – they agonize about difficulty of getting failure-to-repackage papers accepted
  – we hardly ever try to do such work
  – they’re a much older field
  – write/paper/right our power hierarchies that be less entrenched...
  – they are more open profile
  – we don’t have any research results appear regularly in major newspapers/magazines
  – they have rich fabric of blogs as major drivers of discussion
  – crossing traditional power hierarchies
  – we have fewer active bloggers
  – replication crisis was focus of BELIV 2018 workshop at IEEE VIS

Visual Design Process In Depth: Data Sketches
• detailed process notes, from sketching through coding
• Shirley Wu and Nadieh Brehmer

http://www.datasketch.es/