Telling the statistics behind the stories and the stories behind the statistics NISS Graduate Student Research Conference: Saturday, 12 June, 13:30 – 14:30 pm

A. John Bailer <u>baileraj@MiamiOH.edu</u> @john_bailer @statsandstories



Outline

- -1. Disclaimer: Presentation styles are personal and will evolve
- 0. History: my first teaching experience, my first JSM
- NISS GSRC feedback or why my slides changed yesterday
 { changing focus to more general audience ... }
- 2. What should your audience be asking you?
- 3. What statistical concepts to you want to communicate? { a model for a different media for presentation ... }
- 4. Case study: Podcasting with journalists



Feedback from NISS GSRC participants

- Engaging audience: Framing work to be interesting to intended audience ... scaffolding of explanations and examples, keeping the interest of audience ... all attendees get the content of the talk
- Content: How much detail to present ... right amount of technical material ... Presenting a lot of research work in 5-6 slides



Feedback from NISS GSRC participants (ctd.)

- Time management: ... timing keeping audience interest ... pacing ... 'tend to go over time' ...
- Audience questions: Understand material well enough to handle questions ... answering effectively ... answering correctly ... ok to respond offline?
- Slides: simple, clear effective conveyors of message ...



REACTIONS to Feedback ... [1]

- Engaging audience: Framing work to be interesting to intended audience ... scaffolding of explanations and examples, keeping the interest of audience ... all attendees get the content of the talk
- RESPONSE: Data that inspired work and research question to be answered > Methods used so far (lit review) > Opportunities to improve > Your proposed solution > Performance of your proposal > Your solution and the motivating data > Review > To learn more



REACTIONS to Feedback ... [1]

- Engaging audience: Framing work to be interesting to intended audience ... scaffolding of explanations and examples, keeping the interest of audience ... all attendees get the content of the talk
- RESPONSE (continued):
- What do you want the audience to take away from your talk? What were features of talks that you thought were effective? What did your work contribute? Why should the audience care?



REACTIONS to Feedback ... [2]

- Content: How much detail to present ... right amount of technical material ... Presenting a lot of research work in 5-6 slides
- RESPONSE: Depends on the talk and the audience. A job talk will need to show some technical depth and potential for future work (and you have 45-55 minutes!) vs. a contributed paper at a conference (12-15 min). Better to motivate and provide foundation for attendees to build upon.



REACTIONS to Feedback ... [3]

- Content: Time management: ... timing keeping audience interest ... pacing ... 'tend to go over time' ...
- RESPONSE: Practice! Give your talk to an empty room where you talk through your slides to see where transitions are rough; give your talk to collaborators and colleagues for feedback; don't over-practice (relationship between arousal (pressure) and performance for Yerkes-Dodson Law (1908))
 Likely to talk faster in conference / public presentation



REACTIONS to Feedback ... [3]



https://delphis.org.uk/peakperformance/stress-and-thepressure-performance-curve/ (not a distribution but nonmonotonic doseresponse pattern)

0



REACTIONS to Feedback ... [4]

- Audience questions: Understand material well enough to handle questions ... answering effectively ... answering correctly ... ok to respond offline?
- RESPONSE: You are likely to know more than anyone in the audience about the topic. People usually aren't jerks. Better to acknowledge and respond later than rush a response. Applies to classroom teaching as well as presentations.



REACTIONS to Feedback ... [5]

- Slides: simple, clear effective conveyors of talk message ...
- RESPONSE: Content text vs. figures –
- "Aquatic toxicity among C. dubia exposed to phototoxic chemicals from motorized watercraft in a Western Alpine lake"



REACTIONS to Feedback ... [5]

Slides: simple, clear effective conveyors of talk message ... RESPONSE:



Green Pine Trees and Lake (pexels.com)



Man Riding on Green Personal Watercraft on Body of Water (pexels.com)



species.wikimedia.org



REACTIONS to Feedback ... [5]

- Slides: simple, clear effective conveyors of talk message ...
 RESPONSE: Mechanics: Fonts / line spacing + lines per page / color contrast readable from the back of the room.
- ^{8 point < 12 point < 18 point < 24 point < 32 point}
- Black font on white background
- White font on black background
- Yellow font on white background

{ bring copies of slides? List of refs/links to your github repo or web page? }



Next ... Summary (TL; DL <-- presentation analog to

- *** presenting more applied research to a general audience
- S1: Know you audience and tailor your message to them
- S2: Don't fall in love with point estimates no change may be the story
- S3: There are other measures of center that might be more interesting than the center (mean)
- S4: Working with journalists will teach you to target your message



What should your audience be asking you?

Who is your audience?

Is your audience statistically literate?

- Statistical literacy is portrayed as the ability to interpret, critically evaluate, and communicate about statistical information and messages. (Gal 2002 ISR)
- "statistical literacy" refers broadly to two interrelated components, primarily
- (a) people's ability to interpret and critically evaluate statistical information, datarelated arguments, or stochastic phenomena, which they may encounter in diverse contexts, and when relevant
- (b) their ability to discuss or communicate their reactions to such statistical information, such as their understanding of the meaning of the information, their opinions about the implications of this information, or their concerns regarding the acceptability of given conclusions.

Do we have a responsibility to help develop the statistical literacy skills of our audience?



What can we learn from journalists?

"The job of the journalist is to <mark>make the significant interesting</mark>." -Bill Kovach and Tom Rosenstiel's The Elements of Journalism.

Nick Kristoff: Reporters should not lead their reports with a lot of numbers and data but rather tell a story that grabs the reader and illustrates the big data to come later in the story. (RC)

Journalists are supposed to put facts into context [and this] is the foundation of my work as a reporter and my work in the classroom as I teach journalism students how to write news. (R. Pennington)

What is the job of the statistician?



1. Statistician's job?

To conduct and to communicate the results of analyses to have impact.

"Grabbing the reader [audience]" - making the "significant interesting" – "putting facts in context" all seem like important for producing effective and impactful analyses

How often do we think about this?



Working with Journalists – News & Numbers

- My early conversion to thinking about journalism and statistics Cohn and Cope *News & Numbers* classic book for journalists.
- Included questions journalists should ask about research AND statisticians should be ready to answer!
- When teaching a graduate course for non-statistics majors many years ago, I started to include questions inspired by *News & Numbers*



1. Working with Journalists – journalistic challenge?

Example Questions (from Chapter - *Testing the Evidence*):

- How do you know?
- Have claims been subjected to any studies or experiments?
- If studies have been done, were they acceptable ones, by general agreement?
- Results fairly consistent with those from related studies, and with general knowledge in the field? { JB: change in context of time series }
- Finding resulted in a consensus among other experts in the filed?
- Conclusions backed by believable statistical evidence? Degree of certainty or uncertainty?
- Reasonable theoretical plausibility to the findings?



1. Other perspectives: Tim Harford- Financial Times index card

https://medium.com/financial-times/tim-harfords-guide-to-statistics-in-a-misleading-age-652a597c1d88

- 1. Observe your feelings gut reaction
- 2. Understand the claim
- What does it mean?
- Causal?
- What's being left out
- 3. Get the backstory
- 4. Put things in perspective
- Is that a big number
- What is the historical trend?
- Beware 'statistical significance'
- 5. Embrace imprecision
- 6. Be curious
- Go another click
- Treat surprises as a mystery



1. Other perspectives: David Spiegelhalter *The Art of Statistics*

10 Rules of Effective Statistical Practice

- 1. Statistical methods should enable data to answer scientific questions
- 2. Signals always come with noise
- 3. Plan ahead, really ahead
- 4. Worry about data quality *everything rests on the data*
- 5. Statistical analysis is more than a set of computations
- 6. Keep it simple main communication should be as basic as possible
- 7. Provide of assessments of variability with the warning that margins of error are generally bigger than claimed
- 8. Check your assumptions
- 9. When possible, replicate!
- 10. Make your analysis reproducible



1. Other perspectives: David Spiegelhalter *The Art of Statistics*

10 Questions to ask when confronted by a claim based on statistical evidence – trustworthiness of numbers (1-3), source (4-6), interpretation (7-10)

- **1.** How rigorously has the study been done?
- 2. What is the statistical uncertainty / confidence in the findings?
- 3. Is the summary appropriate
- 4. Reliability of the source?
- 5. Is the story being spun?
- 6. What am I not being told?
- 7. How does the claim fit with what else is known?
- 8. What's the claimed explanation for whatever has been seen?
- 9. How relevant is the story to the audience?
- 10. Is the claimed effect important?



1. My personal working list

- 1. What is claimed? and is it appropriate?
- 2. Who is claiming this?
- 3. Why is it claimed?
- 4. Is this a good measure of impact?
- 5. How is the claim supported?
- What evidence is reported?
- What is the quality / strength of the evidence?
- 6. Is the claim reasonable in itself?
- Does prior belief impact my belief? Confirmation bias?
- 7. How does this claim fit with what is already known?
- 8. How much does this matter for me?
- Comparison of population perspective vs. individual perspective?
- Will I change my behavior as a consequence of this?



Ex: World's population could swell to 10.9 billion by 2100, U.N. report finds (NBC News -

https://www.nbcnews.com/mach/science/world-s-population-could-swell-10-9-billion-2100-u-ncna1017791)

What is claimed? and is it appropriate? World pop'n: 10.9B by 2100

Who is claiming this?

- UN Dept of Econ. and Social Affairs World Pop'n Predictions
- Based on "The population estimates and projections contained in this revision cover a 150-year time horizon, which can be subdivided into estimates (1950-2020) and projections (2020-2100)."

Why is it claimed? Analysis and modeling of population counts, fertility, mortality, migration

Is this a good measure of impact?



EXAMPLE: World's population could swell to 10.9 billion by 2100, U.N. report finds

NRC Nowe

ttps://www.nbcnews.com/mach/science/world-s-population-could-swell-10-9-billion-2100-u-ncha1017791

How is the claim supported?

- What evidence is reported?
- What is the quality / strength of the evidence?

Is the claim reasonable in itself?

- Does prior belief impact my belief? Confirmation bias?
- 1. How does this claim fit with what is already known?
- 2. How much does this matter for me?
- Comparison of population perspective vs. individual perspective?
- Will I change my behavior as a consequence of this?



2. What statistical concepts to you want to communicate?

- 1. What did you measure and how did you measure it?
- 2. How did you generate the the data?
- 3. How did you summarize the data?
- 4. What conclusions do you draw from the data?



3. Concepts 3.1 Causality Suggested in headline ... story ... original source

Reverse causality possible

3. Concepts (continued)
3.2 Variables
... Measurement levels
... Validity / Reliability

(often not considered in stat courses)

3. Concepts (continued)3.3 Uncertainty / Variability

... Uncertainty := property of humans [epistemic]

... Variability := trait of system [aleatory]

3. Concepts (continued) 3.4 Data Generation / Producing Data (Tip of the hat to Moore and McCabe) ... Samples (probability vs. convenience) ... Experiments ... Observational * cross-sectional / case-control / longitudinal

3. Concepts (continued)

- 3.5 Describing Data
- ... Distribution
- * location / spread / shape (skewness, modality) ... Relationships
- * numeric variables / categorical variables ... Risk Measures
- * Baseline response rates reported? / probability
- * Relative / Absolute** OR, RR, RD

3. Concepts (continued)

- 3.6 Drawing Conclusions from data
- ... Estimates
- * point / interval
- ... Hypothesis Tests
- * P-values
- * Significance Testing ... Modeling

Acknowledgments

Inspirations: Victor Cohn, Joel Best, David Spiegelhalter, Alberto Cairo

Richard Campbell – partner and co-founder of S+S

Rosemary Pennington – current moderator





4. S+S: History and Context

CAS Quantitative Literacy (QL) considerations – QL FLC

- Challenge to demonstrate infusing QL in Humanities course
- Team teach "News & Numbers" Honors course with Richard Campbell (2009)
- Serving on ASA Board of Directors (2011-13), ISI Council (2009-13), ISI EC PLUS
 - First World Statistics Day (20 Oct 2010)



S+S: History and Context (continued)

- Desire: help broader community understanding the statistical aspects of complicated stories
- Intended audience? Journalists, statisticians, general public, students, teachers (of introductory statistics and more!)
- Inspirations? Freakonomics (economist + writer), Science Friday, NPR Car talk {Aspirational goal: syndication? }
- Model? Nightline (3-5 minute introduction followed by conversation)



25

S+S: Early Decisions

From Start to First Episode:

- Production and recording logistics (radio studio); theme music (stock20.com – play intro – changed with new moderator); web site (purchased domains statsandstories.net; *.biz; *.org)
- Who is "on air"? Moderator + 2 panelists (stat, journalist) + guest
- Topics for Episodes? Mix of serious and fun. Timely topics good.
- Post-production mp3, web site, iTunes, promotion (Twitter, Email)



S+S: Early Decisions– live guests in studio ...





S+S: Early decisions – early review and need to change

Feb 2013 - First guest Tommy Wright (Census) invited + show recorded. Used scripted program notes.

March 2013 – Sent out program for comments ... [+] nice idea / Bob's voice / person-on-the-street [-] too long to get to interview / tech jargon / comments-responses <1.5min ['ear fatigue']

Suggestions:

Point person to coordinate flow of conversation & challenge jargon

Lots of editing of current program + modifications for next recording Mar 2013 – Jim Albert recording



S+S: Early decisions – post-recording issues

- Transcripts before release (accessibility) student workers at first (now pay for transcription).
- Post episode to web page and Apple Podcasts.
- Send out tweet (from Twitter handle @statsandstories feel free to follow! Retweets and favorites help)
- Broadcast email to interested colleagues (ASA, ISI, others)
- Can't track iTunes subscriptions but Google Analytics on www.statsandstories.net



S+S Checklist for Episode

IDENTIFYING and PREPARING for guest

Request biosketch from guest + consider title options + questions for person on the street

BEFORE guest arrives { modify for virtual guests } travel and lodging arrangements for guest PACKAGE production* PERSON ON THE STREET production*

DURING guest visit and recording

Release form signature to guest

Quick intro to recording studio (e.g. position of mics, water caps)

AFTER RECORDING but before release

Combine package + person-on-street segments + guest conversation Transcript



S+S Checklist for Episode (continued)

RELEASE

- Email contacts at Miami, ASA and beyond with announcement
- Twitter (S+S): tweet episode release
- Facebook (S+S, other): announce episode release
- Social media from CAS?
- Social media from University?



Going live! (2013-middle of 2018)





S+S: Evolution- to present form

Partnership with American Statistical Association (ASA) – President Barry Nussbaum Benefits:

- Budget for transcriptions + Hiring part-time producer!
- Connection with speakers introduced by a professional society
- Promotion of program by ASA (ISI also promotes!)
- Connection to ASA-RSS *Significance* editor joint promotion of authors/guests Challenge:
- more frequent and regular release of episodes: move from every 6 weeks to 2 weeks!

Lessons from the professionals? Science Friday visit – regular release and website



Connected to public radio in the US (WYSO)





3. Evolution- to present form

Visit <u>www.statsandstories.net</u> and you will see:

- A carousel of our latest episodes (images: pexels.com)
- A list of topics that we cover
- A back catalog of all of our previous episodes
- Connection to our Soundcloud and Twitter





3. Evolution-mini rebrand

We are able to more modernize our output

- Changed our logo
- Moved the backlog to SoundCloud
- Using our Twitter account in a more engaging way





S+S: Evolution- topics covered?

Broad range of topics and guests ...

- Significance featured authors 53 (P. Guttorp), 63 (N. Thieme), 59 (L. Bornn), 60/66 (A. Georgiou), 91 (A. Carriquiry)
- Sports 1, 18, 21, ... 59, 80
- Health 3, 5, 8, 9, 15, 26, 27
- Politics 29, 51, 66, ..., 90, 92
- Journalists 12, 16, 17, 18, 19, 29, ... 61, 71, 81
- Statisticians 1, 2, 4, 8, 9, 11, 13, 15, 19, 20, 21, 26, 27, 31, 32, 35, 36, 37, 44, 45, 46, 48, 52, 53, 54, 59, 62, 65, 66, 68
- Miami Guests 3, 6, 22, 23, 24, 25



17

S+S: Evolution– changes from start

- New moderator (Bob Long retired) Welcome Rosemary Pennington (public radio and podcast experience!)
- Partnership with ASA (part of ASA presidential initiatives)
 - Funding to support transcription
 - Increased frequency of program release
 - (S+SS introduced)
 - Promotion (Twitter+)



48

S+S: Evolution– frequency of episodes

 With more funding and growth we were able to up the amount of episodes and increase both audience participation as well as book more high profile guests



S+S: Evolution- current state

- All guests virtual (audio quality still good interaction tougher) – good thing we started before the pandemic!
- Target: One episode per week
- External transcription services and hiring a part-time producer



S+S: Evolution— is anyone listening?

- Since our mini rebrand, we switched to hosting to the podcast on Soundcloud which gives us more data on who is listening
- ASA partnership led to ASA promotion earlier last year, these are our numbers from the past month (orange bars are the most listened to single episode on that day, grey bars are our back catalog on those days)



3. Evolution— is anyone listening?

Which Episodes Are Doing Well? (count of >2 minutes listening) – Episode #191 dropped 02 June 2021 – typical listens on recent episodes: 1400-1600



• All time



Explaining Bayes Better | Stats + Stori... 1,276



What Does a National Statistics Office...



1,127

52

Understanding Data in the Digital Ag...



S+S: Evolution— is anyone listening?

Listened to in 148 countries (from Soundcloud – since episode #64).

Here's a quick look at the top ten countries that are listening to Stats and Stories



S+S: Future

Goal remains ... Promoting good reporting and good statistical thinking

Dreaming big ... S+S (2041 followers, <\$25K/year budget) ... SciFri (768K followers, \$20M budget)

ISI partnerships in the future? New Special Interest Group? SIG on Communicating Statistics and Data Science?



S+S: Future

- Context:
 - Climate where opinion viewed as fact ... Echo chamber of news sourcing ... Addressing assertions of Fake news
- S+S Short term:
 - Revenue to support Full-time producer corporate sponsors?
 - Identify other partners
 - Next listener competition?
- S+S Long term
 - Expanding the reach of S+S syndication with NPR
 - Partnering with other efforts Knight proposal *Report for America*



55

Contact information / Questions?

Contact information:

John Bailer

Email: <u>baileraj@miamioh.edu</u> URL: http://www.users.miamioh.edu/baileraj @john_bailer @statsandstories

