Presenting empirical research
Goals

• Enough info to be replicable
• Enough info for results to be convincing
  – My mom says it’s great!
• Limitations: get out ahead of the reader
  – Ignoring doesn’t work
  – All empirical studies have limits!
  – Explain why these limits are reasonable for this study, in this context
Key items

• Methods
  – Data collection
  – Data analysis

• Results

• Limitations
Human subjects: methods outline (approximate)

• 3.0 – high-level overview
• 3.1 recruitment
  – Or last after detailed walkthrough
• 3.2 definition of conditions (if complex)
• 3.2 detailed study walkthrough
  – Might be multiple subsections if complicated
• 3.3 optional collection info
  – If it’s interesting/non-obvious, like you had to instrument something
Methods outline (approximate)

- 3.4 Analysis
- 3.5 Limitations
  - Can also go with discussion
  - I like them upfront to set expectations and avoid “how did they not notice?” questions
Results vs. Methods

• Methods are reproducible
• Dates, counts, descriptives (demographics) go in results later
Methods: Collection: Human subj.

• How did you recruit?
  – Flyers / Mturk / snowball / etc.
  – Were they primed?! / Recruiting message
  – Why this approach?

• What did you pay?

• Ethics compliance
Methods: Collection: Study

• What were the tasks/questions?
  – Include exact protocol as an appendix if possible
  – Was anything drawn from prior work?

• How were participants assigned to conditions
  – Random, round-robin, blocking?

• Any ordering stuff (randomization, alternate)

• How long did it take to participate? (avg, range?)
  – Maybe goes in results?

• Point out decisions that strengthen validity
Data collection, not humans

• Enough info to replicate
  – Hardware used, software versions, network info
Existing data sets

• If using an existing data set, tell me about it!
  – Human or otherwise
  – Don’t make me look up the prior paper
  – Need most of the same info in order to find this credible!
Human subjects: methods outline (approximate)

• 3.0 – high-level overview
• 3.1 recruitment
  – Or last after detailed walkthrough
• 3.2 definition of conditions (if complex)
• 3.2 detailed study walkthrough
  – Might be multiple subsections if complicated
• 3.3 optional collection info
  – If it’s interesting/non-obvious, like you had to instrument something
Methods outline (approximate)

• 3.4 Analysis

• 3.5 Limitations
  – Can also go with discussion
  – I like them upfront to set expectations and avoid “how did they not notice?” questions
Human subjects: methods outline (approximate)

• 3.0 – high-level overview

• 3.1 recruitment
  – Or last after detailed walkthrough

• 3.2 definition of conditions (if complex)

• 3.2 detailed study walkthrough
  – Might be multiple subsections if complicated

• 3.3 optional collection info
  – If it’s interesting/non-obvious, like you had to instrument something
Methods outline (approximate)

• 3.4 Analysis

• 3.5 Limitations
  – Can also go with discussion
  – I like them upfront to set expectations and avoid “how did they not notice?” questions
Methods: Analysis

• Put it here to avoid repeating yourself during results
  – If you do something different in every section, can save for results instead. But unusual that that’s a good idea
Methods: Analysis: Qualitative

• What approach to coding?
• How many researchers / independently
• Inter-rater agreement
• Resolution of conflicts
• (Although qualitative, report some counts of codes for context) … maybe
Methods: Analysis: Quantitative

- Define your metrics (e.g., password strength, earth-mover distance, etc.)
  - And why they are reasonable
- Define your hypothesis tests
  - Why is it appropriate
  - What assumptions had to be checked, potentially
  - A priori power
  - Planned comparisons
  - Post-hoc correction where applicable
  - If complicated, guide to interpret (e.g. logistic reg.)
RESULTS
Overall tips

• Organize by research question

• Avoid wall of stats and numbers
  – Topic sentences, high-level takeaways
  – That are then supported by various metrics/tests
  – *Interpret* statistical results for the reader. What does this result “prove”? Is this meaningful?
  – The stat is not the point, it is supporting evidence for the point!
Start w/ basic descriptives

• People:
  – How many (per condition), demographics
  – Qualitative / small sample: demog table w/ details
    • Use P1 – PX or similar / use IDs based on condition
  – Larger sample, overview table
    • Averages, ranges, quartiles? Compare to census?
  – Consider hypothesis tests to compare conditions
    • Condition 1 is not significantly older, more male than cond 2 …

• Date when data was collected
Further general descriptives

• (optional as own section; might go into results subsections)

• Total items/records/etc.

• Some distribution data
Reporting numbers

• For larger samples, report both number and percent: 49 people (28.2%) or vice versa

• For small samples, avoid percentages as misleading, e.g. 4/5 people vs. 80%
Reporting hypothesis tests

• Report descriptive answer, e.g. condition 2 had mean of 35, condition 1 had mean of 45

• “This difference was significant (T/X2 = xxx, p=0.001)
  – Report p-vals to 3 decimals, or else p < 0.001
  – NEVER say p = 0.000
  – Mention when corrected

• Report effect size (via measure or by using descriptives)
Readable tables

• Use consistent decimal places
• Indicate significant comparisons via asterisk, bolding, etc.
  – This can get quite elaborate
Descriptive graphs

• Plots with error bars (e.g., 95% CI)
• Boxplots and how to read them
  – Band is median
  – Box extends to Q1 and Q3
  – Whiskers vary; most common is most extreme point within 1.5IQR of box in either direction
  – Data beyond whiskers = outlier points
• Stacked bars for Likerts
Choosing graphs

• Choose graphs that illustrate the point: e.g., illustrate a difference that is significant or show two things that aren’t significant and look similar
  – Multi-variate/dimensionality

• If necessary, annotate significant vs. not
Readable graphs

• Default graphs from e.g. R are usually not
• Not too small, not too many things
• Distinguishable colors/shapes
• Clearly labeled axes
• Interpretive captions
LIMITATIONS
Overall goal

• Make it clear to reviewer you know about them
• Explain why they were unavoidable / the best available tradeoff
• Explain what you did to mitigate impact
“Similar to other studies”

• Sampling / representativeness
• Self-reporting issues
• Online study issues
• Various general validity concerns

• Mitigations: pilot/pre-tests, priming, blocking, attention checks, motivations, etc.
• (Generic would apply in any case; prove you designed with them in mind)
“Specific to this study”

• What did you forget to think about (always sthg)
• What is hard in your setting
  – Deception
  – Ecological validity
  – Precision of measure
  – Etc. etc.
Mitigations

• “applies across all conditions so comparisons are valid”
• Better (or not worse) than alternative X
• “A field observation would provide rich data but would not allow controlled experiments/causal analysis” (vice versa)
ADJUSTING FOR SPACE/TIME
Presentations/summaries

• Don’t have enough time for all, what to cut?
• Depends on audience, time (of course), but some ideas:
Highlight main results

• For an audience that might not care a lot about methods
  – But make sure you clarify limitations in interpretation/generalizability so you don’t mislead
Topic audience

• Enough methods to convince of rigor
  – “a standard HCI technique”

• Sample size

• Details of protocol to make tasks clear

• Indicate what is significant, but maybe not details of test, no p-values
Methods audience

• Methods at least equal in size to results
• Details of collection, details of analysis
• High-level results w/ example evidence