INTERACTIONS BETWEEN CONCEPTUALIZATIONS OF CONFIDENCE LEVEL AND ERROR IN THE CONTEXT OF CONFIDENCE INTERVALS

Susan Lloyd (Bucknell University)

eCOTS 2026

Contributors/Collaborators:
Matthew Beckman, Neil Hatfield, and Nicole Lazar (The Pennsylvania State University)

WHY STUDY INTERACTIONS?

- Pilot an approach to educational assessment that allows for the estimation of interaction effects among factors likely to influence understanding related to a construct of interest
- Why study interaction effects?
  - Conventional practices in assessment development often resemble a one-factor-at-a-time design[1, 2, 6]
    - Each item assesses one learning outcome, in isolation, independent of the other learning outcomes
  - Cognitive theory suggests individuals tend to coordinate multiple ways of thinking[4, 5]
  - Instructors are encouraged to foster connected understanding across statistical concepts[3]
- Attributes of confidence intervals studied: *confidence level*, *error*, sample size, p-value, method

1. delMas et al., 2007; 2. Sabbag et al., 2015; 3. Schau & Mattern, 1997; 4. Thompson, 2008; 5. Thompson et al., 2014; 6. Ziegler, 2014

QUALITATIVE DATA COLLECTION

- 6 think-aloud interviews, recruited from upper-level statistics courses
- Students read/discussed their thinking as they engaged with items (probed when necessary)
- Modeled student conceptualizations of confidence level and error using grounded theory[7]

- Preliminary evidence and patterns of interactions between students' conceptualizations of confidence level and error in the context of confidence intervals

7. Strauss & Corbin, 1990

STUDENT 1

Task: Critique an erroneous interpretation of the margin of error for a 99% bootstrap confidence interval as the range of values above (and below) the provided sample statistic.

Student Excerpt: “Usually, when you interpret a confidence interval, it’s important too to mention what the confidence interval is, because if you’re just saying what’s in the quotes [the erroneous interpretation of the margin of error] without the context of it being 99%, they won’t really understand the confidence level, and they need context of that to find out how you got to your answer. Also, I think it would be good to specify the values, because it says the range of values above and below the sample correlation is [sample statistic value]. You could also phrase it as like plus or minus [margin of error value]"

Observations of how this student coordinates their meanings (evidence of interaction):

- Uses meaning for confidence level as a percentage; meaning for error as one of the components involved in a formulaic calculation; meaning for confidence interval as a range
- Constructs a meaning for the margin of error that is akin to the rote formation of a confidence interval

STUDENT 2

Task: Consider the tradeoffs for deciding whether to use a 90% or a 95% confidence interval for inference.

Student Excerpt: "Okay, so a 90% interval would cover less than a 95% interval … but then, also, I think if you did a 95% … yeah a 95% confidence interval would give you a larger interval … But I think if you had a larger interval [discussing this as a result of a higher confidence level] and more data to take from, I think maybe you'd have less variation and a more accurate answer. And then also your significance level would probably go down, so then .. it would narrow the .. not the requirements, but like the bar. … The window that you have to be in. So then that could also make it more ... Know that your answer's more accurate"

Observations of how this student coordinates their meanings (evidence of interaction):

- Uses meaning for confidence interval as covering the cases; meaning for confidence level in terms of its normatively accepted relationships with significance level and interval width; meaning for error as variation
- Constructs a meaning for error that no longer has the formulaic aspect previously mentioned
- Constructs a meaning for error that is more in-line with the precision fallacy in the context of error

STUDENT 3

Task: Identify a true statement about a 99% confidence interval for the correlation that ranges from a negative value to a positive value.

Student Excerpt: "[I]t's [the margin of error is] not the total range, it's the … radius ... from that analogy. … but (d) says that decreasing the confidence level to 95% would result in a wider confidence interval. That's not true, it would be the opposite relationship, decreasing the confidence level to 95% would result in a narrower confidence interval .. because .. we're less confident that it would fall within a narrower interval."

Observations of how this student coordinates their meanings (evidence of interaction):

- Uses meaning for the formulaic relationship between the confidence level and interval width; meaning for the margin of error as the radius of the interval
- Constructs a meaning that no longer aligns with a productive interpretation of confidence intervals (places the confidence in a particular interval, rather than in the method)

STUDENT 4

Task: How the width of a 90% confidence interval compares to the width of a 95% confidence interval (for the same sample).

Student Excerpt: "Okay. So compared to the 95% confidence interval, the 90% confidence interval … I'm pretty sure it's going to be wider because since there's less confidence like a larger or no, no the 90. Yeah, it's going to be wider because we're more confident in it. That means that there's probably going to be a larger range, so we want more room for error .. So [the 95 would be] wider because it has a larger standard error since the researcher is less [misspeak – should be more] confident."

Observations of how this student coordinates their meanings (evidence of interaction):

- Uses meaning for confidence level as a success rate; meaning for error as a measure of efficiency
- Constructs a meaning for confidence level that is no longer tied to the multiplier
- Constructs a meaning for confidence level that is closely intertwined with error (the confidence level, not the sample, determines the value of the standard error)

STUDENT 5

Task: Address the relationship between the confidence *level* and the confidence *interval*.

Student Excerpt: "[A]s you decrease the confidence level, the confidence interval width will also decrease. And that's because like before you're going to want, you're going to accept more error so that's going to allow for you to be wrong more times. Like if it was an 80% confidence interval, you're saying that in only 80% of the confidence intervals you make the sample mean will be included, so that's going to be a lot more narrower, because you have a higher chance of the true level not being in the confidence interval."

Observations of how this student coordinates their meanings (evidence of interaction):

- Uses meaning for confidence level in terms of its effect on interval width; meaning for error as the interval having the property of making "errors" or mistakes
- Constructs a meaning for the confidence level as the percentage of confidence intervals that are expected to capture the sample statistic
- Articulates a relationship between confidence level and error to reconcile their existing meanings

STUDENT 6

Task: Estimate the standard error from the provided bootstrap distribution.

Student Excerpt: "[T]he standard deviation *of* that sample statistic is going to be … [5 second pause] … okay hold on .. thinking another way. … The 95% confidence interval is two standard deviations away in each direction … Okay so it's two standard deviations away. Which means … that the .25 [misspeak – should be .025] is 2 standard deviations away. .. Hold on, no, the .25 [misspeak – should be .025] is outside of it. … Oh … [5 second pause] … oh well then shouldn't point .. the standard error just be .. twice what the .. outside of the confidence interval is, so the standard error is going to be .025 times 2"

Observations of how this student coordinates their meanings (evidence of interaction):

- Uses meaning for confidence level as related to the decision threshold for rejecting the null hypothesis; (corrected) meaning for standard error as the standard deviation of the sample statistic
- Constructs a meaning that is consistent with an apparent association between 95% confidence and two standard deviations in each direction
- Constructs a meaning that equates the standard error with the rejection region

SUMMARY

- Instances in which each student coordinates meanings for confidence level and error
  - Importance of fostering connected understanding across statistical concepts

- Support for incorporating interaction effects into the design of educational assessments
  - Allow instructors to achieve greater clarity on how students coordinate meanings for attributes
  - Encourage students to use ways of thinking about one concept to enhance understanding of another concept

- Other takeaways: prevalence of lexical ambiguity in statistical terms; students apply understanding from everyday language
  - “error”: the interval has the property of making mistakes
  - “range”: confidence interval as a range of values
  - “confidence”: synonymous to probability

REFERENCES

1. delMas, R., Garfield, J., Ooms, A., & Chance, B. (2007). Assessing students' conceptual understanding after a first course in statistics. *Statistics Education Research Journal, 6*(2), 28-58.
2. Sabbag, A., Garfield, J., & Zieffler, A. (2015). Quality assessments in statistics education: A focus on the GOALS instrument. In M. A. Sorto (Ed.), *Advances in statistics education: Developments, experiences, and assessments, IASE 2015 satellite conference proceedings*. International Association for Statistical Education.
3. Schau, C., & Mattern, N. (1997). Assessing students' connected understanding of statistical relationships. In I. Gal & J. B. Garfield (Eds.), *The assessment challenge in statistics education* (pp. 91-104). IOS Press.
4. Thompson, P. W. (2008). Conceptual analysis of mathematical ideas: Some spadework at the foundation of mathematics education. In O. Figueras, J. L. Cortina, S. Alatorre, T. Rojano, & A. Sepúlveda (Eds.), *Proceedings of the Joint Meeting of PME 32 and PME-NA XXX* (Vol. 1., pp. 31-49). Psychology of Mathematics Education.
5. Thompson, P. W., Carlson, M. P., Byerley, C., & Hatfield, N. (2014). Schemes for thinking with magnitudes: A hypothesis about foundational reasoning abilities in algebra. In L. P. Steffe, K. C. Moore, L. L. Hatfield, & S. Belbase (Eds.), *Epistemic algebraic students: Emerging models of students' algebraic knowing, WISDOM^e Monographs* (Vol. 4., pp. 1-24). Laramie, WY: University of Wyoming.
6. Ziegler, L. A. (2014). *Reconceptualizing statistical literacy: Developing an assessment for the modern introductory statistics course* [Doctoral dissertation, University of Minnesota]. University Digital Conservancy.
7. Strauss, A., & Corbin, J. (1990). *Basics of qualitative research: Grounded theory procedures and techniques.* Sage Publications.