The Future of Engineering Education
2024 ASEE Annual Conference
June 23 - 26 | Oregon Convention Center, Portland, OR

Paper ID #41422

When Is It Relevant? A Collaborative Autoethnographic Study by Engineering Students on Statistical Variability

Leslie Bostwick, Franklin W. Olin College of Engineering
Alex George, Franklin W. Olin College of Engineering
Trinity Lee, Franklin W. Olin College of Engineering

Trinity Lee is an undergraduate engineering student at Olin College of Engineering pursuing a B.S. in Engineering with a concentration in Computing. She has always been passionate about engineering education and has been part of multiple research labs investigating engineering education and research. At Olin, she has worked with The Kern Entrepreneurial Engineering Network (KEEN) to improve research accessibility to undergraduate students and conducted research on engineering education and its relationship with variability as part of the Uncertainty Research Lab.

Dr. Zachary Del Rosario, Franklin W. Olin College of Engineering

Zachary del Rosario is an Assistant Professor of Engineering and Applied Statistics at Olin College. His goal is to help scientists and engineers reason under uncertainty. Zach uses a toolkit from data science and uncertainty quantification to address a diverse set of problems, including reliable aircraft design and AI-assisted discovery of novel materials.

When Is It Relevant? A Collaborative Autoethnographic Study by Engineering Students on Statistical Variability

Abstract

This research paper presents an exploratory study of engineering students' evolving views of statistical variability. Statistics educators broadly agree that statistical thinking is thinking about variability. However, certain engineering failures and recent literature suggest that variability is neglected in engineering practice—to dangerous effect. To better understand young engineers' perceptions of variability, we conducted a collaborative autoethnographic study with three engineering student participant-researchers. Data collection and analysis ran for ~6 months and followed the quality in qualitative research (Q3) framework to actively promote research quality. Our reflection data illustrate how numerical reporting conventions, formative life experiences, and professional aspirations can all affect a young engineer's perception of the relevance of variability. We conclude with a discussion of implications for instructional practice.

Introduction

Variability—the phenomenon of non-identical values—is core to modern science. The move beyond calculating averages to the study of real variation is one of the most important scientific developments of the 19th century [1]. Ernst Mayr [2] positions variability as fundamental to understanding evolution through "population thinking." Statistics as a discipline exists in large part to develop techniques to study variability: Statistics educators broadly agree that "(a)ny serious discussion of statistical thinking must examine the role of 'variation'" [3]. In this work, we use the phrase "statistical variability" to refer to variability as it is viewed by statisticians: an ubiquitous phenomenon to be directly studied and modeled.

Despite its importance, prior failures in engineering can be traced to the neglect of variability [4]. In the 1950s, design of aircraft for "the average man" led to uncontrollable aircraft [5]. At the height of this issue, as many as 17 pilots crashed in a single day [6]. Similar issues persist in modern engineering practice: As of writing, female passengers are crudely modeled as a scaled version of the male median in automobile crash test practice [7]. The Government Accountability Office [8] asserts that these practices lead to higher odds of female passenger injury, quantified as 47% higher in 1998–2008 by Bose et al. [9].

Furthermore, variability is not a strong focus in engineering education—in research or practice. A systematic review of education literature on mathematics related to engineering found only two out of 5,466 articles that discussed "uncertainty" or "error" [10]. A scoping review of textbooks on engineering course reserve lists found that only 11% of surveyed textbooks mentioned "variability" [11].

If variability is viewed as irrelevant to engineering practice, design issues such as "the average man" and disparities in accident outcomes may persist. In our previous study of practicing engineers, a few (3/24) participants consistently did not factor variability into their design decisions, while the remaining majority (21/24) either acknowledged that variability exists or actively targeted the consequences of variability [4]. These results suggest that many practicing engineers regard variability as important to consider in design, while some practitioners may not yet see its relevance and are hence prone to making dangerous decisions that neglect variability. However, that work used a structured interview approach and did not study relevance *per se*; participants in that study were confronted with variability in the form of multiple observations, which directly signaled the existence of variability, and hence implicitly signaled its relevance. A study of relevance should take place in an *in vivo* setting: For practitioners this is the workplace, while for engineering students this is the classroom.

Hence, variability is often *considered* irrelevant in engineering education, while it *is certainly* relevant to engineering practice. Given the demonstrated importance of variability to engineering safety, an engineering pedagogy that suggests the irrelevance of variability is unacceptable. Research is necessary to determine the ways current engineering pedagogy contributes to young engineers' views on variability, motivating the present study.

Our goal with this work was to investigate young engineers' perception of the relevance of statistical variability to their work, and discover experiences that may promote relevance. We framed this work using concepts from statistics education research and social learning theory. To access rich, personal experiences of variability, we used a collaborative autoethnographic (CAE) approach. Our CAE approach required us to narrow the scope of our investigation to a small number of participants, but enabled a deep investigation of episodes that demonstrate the relevance (or not) of statistical variability within the cultural context of an undergraduate engineering program.

Conceptual and Theoretical Frameworks

This work drew on statistics education for its conceptual framing and used communities of practice as a theoretical framework [12]. Given its importance to engineering, our conceptual focus is on variability. In particular, we drew on research at the intersection of statistical and engineering practice: the NAT Taxonomy describes how engineers either *neglect* variability, *acknowledge* its existence, or make decisions to *target* its consequences [4]. These insights served as sensitizing concepts [13] to call our attention to variability—especially to participants' sense of its existence, relevance, and consequences.

Communities of practice (COP) is a social learning theory developed by Wenger [14]. COPs are a mesoscale unit of analysis: they transcend the individual by considering social processes, but maintain a focus on groups of individuals who mutually engage in a shared practice (such as students in a classroom, or a team of engineers within a larger firm). The "practice" in a COP is thought to be a negotiation of meaning unfolding through dual processes of *participation* and *reification*. Wenger's use of *participation* follows common usage, referring to "the social experience of living in the world in terms of membership and active involvement in social enterprises" (p. 55). *Reification* is a less common term, referring to "treat[ing] (an abstraction) as substantially existing, or as a concrete material object " (p. 58). Examples of reification include concrete artifacts (e.g., paper reports), documented procedures (e.g., rules for writing reports), and codified positions within an organization (e.g., junior vs. senior engineer).

Reification has a double-edge; for instance, reifying practice into a tool can make a sequence of actions effortless, but also has the potential to prevent further developments to practice. As Wenger writes (p. 59),

> "… I want to preserve the connotations of excessive concreteness and projected reality that are suggested by the dictionary definition. Indeed, no abstraction, tool, or symbol actually captures in its form the practices in the context of which it contributes to an experience of meaning."

Given the noted lack of attention to variability in engineering research [10] and education [11], we expected the student participant-researchers in this study to struggle to identify the presence and relevance of variability in their engineering coursework. Reification as "excessive concreteness" is one mechanism by which variability is rendered invisible: If the tools that engineering students use in school neglect variability, this not only signals that variability is not relevant—it implicitly claims that variability does not exist. Adopting COP as a theoretical lens attuned our study to this possibility.

To summarize the framing of this work, we defined our social reality under investigation (SRUI) in terms of its scope (the extent of the study) and nature (our theoretical assumptions) [15].

- Scope: The student participant-researchers' (Alex, Leslie, and Trinity's) individual experiences with variability, particularly its perceived relevance, within the cultural context of their engineering program and developing identities as engineers.
- Nature: The dual processes of participation and reification associated with a community of practice may affect perceived relevance. In particular, we expect that certain reifications in engineering education render variability invisible and suggest irrelevance.

Research questions

We sought to answer two research questions:

RQ 1. To what extent do engineering students perceive statistical variability as being relevant to engineering practice?
RQ 2. What experiences affect their views and how are they situated in the broader cultural context of engineering?

While the scope of our SRUI was considerably more focused than RQ1 suggests, this focus enabled the use of methods that allowed us to deeply investigate individual experience in the context of a broader cultural experience.

Methodology and Methods

This work was determined to be IRB exempt by Brandeis University's IRB and followed a human subjects protection protocol (#23232R-E). Elements of this protocol were designed to promote research quality through the lens of ethical validation [16], described in this section. We used the *quality in qualitative research* (Q3) framework to actively promote the validity and reliability of our work through making and handling of data [16], [17]. This work was part of a larger study on both variability and mathematical modeling in engineering student culture; below, we present an episode from this context to illustrate our ongoing consent procedure.

Collaborative Autoethnography (CAE)

Autoethnography (AE) is a combination of autobiography and ethnography [18]. AE is a qualitative research method that uses study of the self in a particular cultural context as a means to understand the experiences of those in a similar context [19]. Collaborative Autoethnography (CAE) builds upon the strengths of AE by situating reflection in a collaborative context, where additional meaning can emerge through co-construction among participant-researchers [19].

We chose CAE for its affordances to study relevance: The participant-researchers in this study are uniquely positioned to describe how they (as engineering students) experienced and perceived variability in their studies. Autoethnography provided a framework for the individual participant-researchers to study their own experiences—the students directly documented and analyzed when variability was relevant to them, and what experiences influenced their views. The collaborative nature of CAE highlighted differences in the participant-researcher's experiences, leading to a richer study of the SRUI.

Research Team

The research team consisted of three current students and one faculty member, all at Olin College. At the start of this project, Alex, Leslie, and Trinity (the student participant-researchers) had just completed their first-year at Olin College, including the class *Modeling and Simulation of the Physical World* (ModSim) taught by Zach. Alex, Leslie, and Trinity all reached out to Zach with an interest in pursuing research over the summer, and Zach pitched the idea of using a

CAE approach to study their learning experiences: Zach had recently redesigned ModSim to emphasize statistical variability and was interested in how the students responded to the revised course. Alex, Leslie, and Trinity agreed, and we began the study in the Summer of 2023.

Methods

Data collection for this work took place within the context of a broader study of engineering student culture and mathematical analysis—this study focuses on our discussion of variability. Prior to data collection, the team discussed ideas of variability (including the NAT Taxonomy [4]) as sensitizing concepts [20]. The team conducted ~200 hours of data collection and collaborative analysis over 10 weeks during the Summer of 2023. We shared, developed, and referenced a "CAE Methods" document as a living "audit trail" [21] to promote process reliability [17]—this is reported in Appendix A. This methods document outlined a process following the concurrent CAE model of Ngunjiri et al. [19], including independent written reflection and collaborative discussion. In each cycle, the students revised a set of prompts, wrote an individual reflection responding to the prompts, studied each others' reflections separately, then participated in two discussions of their reflections—one with students only, and one including the faculty member. Both the prompts and reflections were guided by the critical incident technique [22] to reduce self-report bias.

In the Fall of 2023, Alex and Trinity carried out additional observations in one of their major-related technical courses.[1] This also followed a CAE approach, but with reflections written in response to additional coursework. Alex and Trinity attended their courses and took usual course notes,[2] but supplemented this with jottings during the course and completed a field notes reflection template after each course session. We consulted an established anthropologist to provide input and training to Alex and Trinity in how to take effective field notes. The team met weekly during the semester to review and discuss their field notes.

Alex, Leslie, and Trinity conducted data analysis through their individual review and discussion of reflections. For each round of reflection, each student participant-researcher reviewed and compared the others' reflections. We analyzed the data collaboratively through discussion: First with the students alone, then in a second meeting with the faculty member, each time taking detailed notes on the discussion. The team discussion notes served as a record of this first-pass analysis of the reflection data, which fed into the next stage of analysis.

Zach assembled the present manuscript by re-reading notes from team discussions, individual reflections, and field notes. He selected three episodes (one from each student) relevant to statistical variability and students' perceptions of relevance (while variability was a focus of this

---

[1] Leslie expressed an interest in similar extended observation, but was on international study abroad at the time. We were unable to navigate the international IRB process in time to support her work.

[2] Instructors of these classes consented to observation, but were not a major focus of this study.

study, data collection had a broader scope). Zach performed initial coding of the selected episodes to build fluency with the data [23], selected student reflections, and connected the episodes to the study framing. He used verbatim student reflections, making only small edits (reordering and trimming) to promote readability. The students reviewed and commented on the manuscript, in order to ensure their voice and perspective was preserved in the report.

Research quality considerations

From the conception of this project, the team focused on addressing the most significant threat to the validity of our work: the power disparity between faculty and students. There is a clear power disparity between faculty and students, which may influence students to simply report what they believe a faculty member expects. Since there was no deception in our research methods, Alex, Leslie, and Trinity were fully aware of Zach's interests in statistical variability. Our use of the critical incident technique [22] was intended to partially address this concern, but we felt that a more comprehensive approach was necessary. This concern was strong enough that Zach seriously considered canceling this study. However, through engagement with the CAE literature and drawing inspiration from other CAE projects [24], the team arrived at a satisfactory framing and procedure to help mitigate the threat of power disparity.

Our framing drew on ideas of relational ethics and ethical validation. Relational ethics refers to ethical considerations in research involving close personal contacts, which commonly arise in (C)AE research [25]. Relational ethics requires researchers to "initiate and maintain conversations" with participants in a dynamic consent process. Ethical validation refers to the empirically-supported observation that actively seeking to do justice to all research stakeholders will lead to higher-quality research outcomes [16]. These ideas guided us to collaboratively design an ongoing consent process (Appendix B), and convinced us that pursuing such an approach would make this study ethically feasible and promote research quality.

We implemented relational ethics in making and handling our data [17]. We collaborated as a team to define mutually-agreeable procedures for both the CAE and reporting phases of the project (see Appendices). While Zach created an initial set of reflection questions, the full team reflected on the relevance of those questions to the students' lives, and the students developed new reflection questions to ensure resonance with their own experiences. This ensured alignment of the reflection data with participant experience, promoting communicative validation [17].

Inspired by [26], we present one "key episode" that illustrates our ongoing consent process. Late in the Summer phase of the project, our discussion considered economic disparities between students, and how this affects their educational experience. While not strictly related to variability, this was an aspect of engineering student culture that was related to the broader study, and clearly important to Alex, Leslie, and Trinity. By chance, Zach was preparing a "pedagogy bootcamp" for new faculty at Olin College, and asked the students for permission to include some of their reflections as background to help new faculty understand the socioeconomic dynamics at the College. Following our reporting guidelines (Appendix B), Zach prepared a set of quotes from a team discussion to use in the bootcamp and sent them to the team for review. Due to the personal nature of the discussion, the students exercised their unilateral veto (Appendix B) to strike a few quotes from the proposed release. In particular, Leslie flagged a quote by Alex as potentially sensitive, which Alex ultimately decided to veto.

This episode illustrates the quality-promoting features of our approach. The case of one student flagging another student's quote illustrates our principle of mutual protection in action (Appendix B). The vetoing of multiple quotes illustrates that our attempt to create a safe space was successful, as the students shared things with the whole team (students and faculty member) that they would not want to share with a general audience. These observations are encouraging evidence of our success in mitigating the power disparity threat to validity.

Results

In this section we present reflections from each of the three student team members. The selected episodes illustrate different themes related to our research questions.

Alex: Different reifications of numerical values

Alex's experiences illustrate how reifications of numerical values codify variability as either ignored or addressed. They detailed some of their earliest experiences,

> (Alex) "The clearest memory I have of using single values that I question now was in high school physics. Physics had a lot of referencing tables that told you the specific heat of different materials. We never questioned these tables or really talked about where these values came from, this was just a fact about the material. They were just tables in our textbooks that we used sticky notes to mark so we could reference them easily. It never even crossed my mind that in reality, a material's specific heat would probably be slightly different based on the manufacturing process and other factors."

Some quantities are commonly reified as single values: This is an ubiquitous feature of engineering student culture, as common pedagogical approaches reify a sole correct answer as a single value [27]. As Alex described, this presentation gives the impression that such values are immutable constants—that variability does not exist. This impression of fixed value may persist until disrupted, as Alex recalled from participating in a class on Thermodynamics,

> (Alex) "In class today we were calculating the heat transfer coefficient. I realized I had assumed this was just a property, that there was no variability / it's constant for a given fluid. In practice problems and homework we were always given the heat transfer coefficient. We had mentioned that there are lots of factors affecting it, but I either didn't

> fully process this or (more likely) assumed that we typically treated it as a property like we do many other things (conductivity and specific heat, for example). Lots of factors go into it, like the velocity of the fluid."

Alex's reflection illustrates how their developing experience of engineering practice influences the sources of variability that they either ignore or address: Their early experience with heat transfer coefficients dealt exclusively with reified single values, much like their high school physics textbook. However, participating in the calculation of a heat transfer coefficient revealed the potential variability in these quantities. While viewing heat transfer coefficients as fixed constants may be sufficient for certain student cultures, full participation in an engineering culture that designs thermodynamic devices (e.g., cooling fins or combustion engines) requires an appreciation of the factors that affect a heat transfer coefficient—sources of variability.

Not all engineering quantities are reified as single values; geometric variability is reified by dimensional tolerances. Alex described their experience with such tolerances,

> (Alex) "I take variability into account when I am doing machining/fabrication. Variability can come from a lot of places in machining, whether it be the material you're using, the tool you're using, the tolerances of the machine, etc. Variability matters because we want the part we are making to actually serve its purpose. We account for variability here by specifying the tolerance that we want our part to be and by measuring our part multiple times. Typically we don't need a part to be the exact correct size, but there is a range that we want. As long as we hit that range we are within tolerance."

Alex clearly found variability relevant in machining practice: They listed several sources of variability and detailed their approach to checking parts for acceptable dimensions—they targeted the consequences of dimensional variability [4]. Here reification reinforces the importance of variability: It provides a range for checking said parts. Note that the form of reification here makes variability impossible to miss—tolerances take the form of a range, which provide an unmistakable signifier of variability. This is an example from engineering practice where variability is regarded as so important that it is a key feature of practice and is reified in documentation as a range.

Leslie: Variability and social norms

Leslie's experiences illustrate a personal encounter with variability from long before ModSim,

> (Leslie) "My first experience with variability was with women's fashion, which is a pet peeve of mine. I have always been tall for my age, so women's pants were never the right length for me. I remember that 'ankle length' pants would stop halfway up my shins. It wasn't until high school that I found a solution: I started wearing vintage men's Levi's,

> where I could get different lengths of jeans that weren't hundreds of dollars. The limitation is that the fast fashion industry has pushed these 'standardized' sizes to produce clothing faster, but women have a range of proportions and sizes so no clothing fits anyone well."

Leslie described experiencing an aspect of her life as an outlier—a tall person whose dimensions sit outside the "standardized" dimensions reified by the fashion industry. Her participation in society conflicted with this reification of standard sizes, as she had to choose between spending an inordinate amount to buy tailored pants or wearing men's clothing to find better-fitting options. Her experiences sensitized her to variability writ large,

> (Leslie) "Variability is kinda everywhere. I feel like I'm always passively taking variability into account, whether I'm using a left handed scissor with my right hand, cooking, or building a set piece for a musical. There is physical variability (pant lengths), theoretical variability (project budgets), and emotional variability (acting a certain way to produce the best outcome). When enough of society's norms don't work for you, you learn to adapt and be adjusting consistently."

Leslie's perspective contains fundamental elements of statistical thinking; namely, recognizing the omnipresence of variability [3], [28]. She credited this perspective to her conflict with society's norms. Leslie's experience also illustrates a key insight for equitable design—designing for inclusivity fundamentally requires targeting variability.

Leslie's experiences also illustrate the COP concept of *multimembership*—the effects of simultaneous membership in multiple COPs. Her participation outside engineering education sensitized her to recognize variability in her engineering education. These experiences work against the features of engineering student culture that emphasize a single correct answer [27]; put differently, features of Leslie's identity promoted transfer of learning about variability.

Trinity: Relevance to a desired COP

Trinity's experiences illustrate a distinction between learning outcomes and perceived relevance. At the early stages of this research project (summer 2023), she expressed a lukewarm reaction towards variability,

> (Trinity) "Variability is an interesting concept that I'm now aware of, but I haven't used it in practice. Understanding something like 'everyone is different' or something like that is pretty basic, but I think putting that into an engineering context in my field hasn't happened yet or isn't as important in my context (software engineering). Everyone's life and learning journey is different so it's not going to be possible to have everyone agree on such things. For example, both me and Alex did not seem to notice variability as

> prominently in our everyday lives compared to Leslie. I suppose that's why research like this exists."

Trinity's reflection demonstrates reflexivity: She compared her perspective with her peers and gave a window into her perceptions of relevance. At this point in time though, variability had not yet entered her practice and she viewed everyday examples of variability as "pretty basic." We note, however, that she achieved the cognitive learning goals associated with ModSim (specifically with respect to variability), and even imagined transferring those skills to her intended major of software engineering,

> (Trinity) "I think that user unpredictability is a valid form of variability. Sure this variability isn't life-threatening—I mean, it's software—but the variability of how different people would interact with a system needs to be accounted for, or else the system may break. Limited outcome possibilities also play a factor. If we assume there are only a few options and that variability doesn't affect the outcome that much, we tend to disregard it or take the average of all the outcomes because the standard deviation is so small and insignificant."

Trinity's perspective includes sophisticated elements of statistical thinking, such as quantifying variability (standard deviation) to justify cases where variability can be safely ignored [4]. She also imagined physical reasons for variability in software engineering. However, at this point in time, she still viewed variability as a minor concern—without important consequences and to be ignored if possible.

It was not until well after summer 2023 that Trinity expressed an unqualified sense of relevance of variability. While collecting field notes in a class on electronic circuits, she latched onto the concept of metastability in logic circuit states. In a group meeting she spoke animatedly about what she had learned and guided us through her field notes,

> (Trinity) "We talked about metastability in states: an example of uncertainty and variability in computer architecture and how engineers account for it. The only way to get out of metastability is through thermal noise. We take into account how, in theory, our circuit diagram would work, but in real life we need to take into account these edge cases [uncertainty & variability] that will occur, and design around it. I liked it since we're talking more about application rather than just theory."

In contrast with the "basic" examples of variability Trinity described above, this example requires a more sophisticated understanding of relevant physical phenomena—metastability. Metastability (in electronic circuits) is the ability of a digital circuit to persist in an undefined state (neither 0 nor 1) for an unbounded length of time [29]. As Trinity detailed in her field notes, metastability is an unavoidable consequence of variability (e.g., in the arrival time of electronic pulses), and understanding this phenomenon requires an appreciation of variability.

This appreciation of variability is nontrivial: Chaney and Molnar [29] wrote "Significant systems failures have resulted from this fundamentally inescapable problem that is generally not appreciated by system designers and users." Such variability continues to confuse even practicing circuit designers: Ginosar [30] reviewed 14 design flaws in synchronization and emphasizes that "(l)ogic validation tools are typically incapable of detecting any errors in such synchronizers." In this case, the reifications used by practitioners (certain logic validation tools) encode erroneous assumptions about the physical world that neglect sources of variability and lead to failed circuit designs.

While teaching this appreciation for variability is difficult, training engineering students to recognize the sources and consequences of variability is crucial to training engineers who can design safe systems: circuits that account for metastability, human interfaces that work for a diversity of people, or thermodynamic systems that are designed for variable inputs.

Limitations and Discussion

Our choice of methods and sample imposed certain limitations: Key among these were a full team awareness of variability as our research focus. This clearly biased the participant-researchers towards noticing variability in their coursework more than they might otherwise—Our study cannot discern how students would "naturally" experience variability. However, this bias also serves as an important backdrop: When our participants found variability surprising, *despite* the clear framing of the study, it underscores the degree to which variability is rendered invisible and irrelevant by present engineering pedagogy.

While our main interest is engineering practice, our sample consists of engineering students, which necessarily limits the conclusions we can draw about engineering practitioners. Studying students is a useful proxy for studying young practitioners, as engineers frequently start their career as recent students. Prior work [4] has shown that some practicing engineers continue to neglect variability later in their career—the results from this study suggest that features of engineering pedagogy may contribute to this behavior. However, future work with direct study of practicing engineers is necessary to fully understand their views on the relevance of variability to engineering practice.

All three student participants described concrete ways in which variability seemed relevant (or not) to engineering practice (RQ 1), and the specific experiences that influenced their views (RQ 2). Prior evidence suggests that variability has not been considered relevant to certain quantities in engineering practice, despite the demonstrated potential for failure: As noted in the introduction, the neglect of variability [4] in automobile [8] and aircraft [5] design has led to injury and death.

For Alex, the relevance of variability was a surprise, as many of the numerical values they had previously encountered in science and engineering courses were presented as reified single values. Variability seemed irrelevant largely because cultural norms in teaching technical subjects implied that variability did not exist. However, after engaging with variability in their engineering training, they recognized the relevance of variability to engineering practice. This perception of relevance was modulated by the particular form of reified values: While Alex was surprised at variability in heat transfer coefficients (presented as a single value), they found variability in part dimensions unsurprising (presented as a nominal value with a range).

Alex's experiences suggest that the neglect of variability in early technical courses may train engineers to expect the *im*possibility of variability in certain numerical quantities. Simply discussing the possibility of variability is unlikely to disrupt this view, as Alex described,

> Alex: "We had mentioned that there are lots of factors affecting it, but I either didn't fully process this or (more likely) assumed that we typically treated it as a property like we do many other things (conductivity and specific heat, for example)."

A simple way to consistently signal the possibility of variability is to report ranges along with nominal values. This is commonly done in engineering drawings, and could easily be implemented in engineering pedagogy.

For Leslie, the relevance of variability was unsurprising, as she had personally experienced the consequences of variability from an early age. Her visceral feelings towards reified clothing dimensions ("a pet peeve of mine") seem to contribute to a statistically sophisticated mindset: an appreciation for the omnipresence of variability [3], [31].

Leslie's experiences with clothing dimensions are similar to the issues encountered in "the average man" debacle [5] and in modern crash test practices [8]: all involve the reification of anthropometric statistical summaries into design choices. While dimensions such as height and waist diameter are correlated, reifying this statistical association into sizes such as S, M, and L requires those who fall outside the trend to buck implicit norms: In Leslie's case she "found a solution: (she) started wearing vintage men's Levi's." Similar to how intersecting marginalized identities can lead to unique experiences of marginalization [32], variation across multiple dimensions can break the reified statistical assumptions embedded in design practices. Designing for *full* inclusion requires grappling with the multidimensional nature of variability.

For Trinity, variability did not seem relevant to her intended major of software engineering—the COP she seeks to join. Despite achieving significant cognitive learning outcomes related to variability—including successfully transferring her learning to software engineering practice—Trinity did not initially regard variability as relevant to her engineering practice. It was only after encountering a striking, discipline-specific instance that Trinity started to view variability as clearly relevant to her engineering practice.

In Trinity's case, the collaborative nature of our methods and the clear study focus on variability underscore the power of forces in STEM training that suggest the irrelevance of variability—despite the clear messaging from the faculty member and the rest of the research team, Trinity did not find variability relevant to her engineering practice for much of the study. As seen in Alex's experience, the reification of single values is one feature of engineering pedagogy that renders variability invisible and hence irrelevant.

Trinity's experiences illustrate both a challenge and an opportunity in teaching statistical thinking about variability. Statistics educators have long bemoaned the difficulty teaching statistics [33] and worried that engineers do not use statistics in practice [34]. Trinity's experiences illustrate a challenge beyond cognition or transfer of learning—a persistent lack of perceived relevance. This is an important challenge, as it suggests that a pedagogy aimed only at cognition or transfer may not succeed in overcoming challenges such as "the average man" debacle [5] or the ongoing inequities in automotive safety [9]. Engineers who can think about variability but doubt its relevance—or its existence—are unlikely to target variability in practice.

An opportunity for teaching statistics is to draw inspiration from Trinity's experiences: Domain-specific examples that clearly signal the importance of variability to an engineering community's practice are powerful demonstrations of relevance. Such a pedagogy will require broad engagement from engineering faculty in their individual courses, but the examples themselves would not demand much course time or instructor effort. As Alex's experiences illustrate, such examples would benefit from going beyond mere discussion to involve legitimate participation. The importance of variability—for both inclusive and safe engineering design—suggests that such course development efforts are worth implementing as future work.

Acknowledgements

Greses Pérez suggested the extension of CAE into the Fall of 2023. Caitrin Lynch lent her expertise as an Anthropologist to help guide the Fall 2023 autoethnographic work. Nicola Sochacka reviewed several versions of this work and provided valuable feedback, particularly on the qualitative methods employed. Two reviewers provided detailed comments that improved the clarity of this work. This material is based upon work supported by the National Science Foundation under Grant No. 2138463.

References

[1] T. M. Porter, *The rise of statistical thinking, 1820-1900*, New edition with a new preface by the author, New paperback. Princeton, New Jersey: Princeton University Press, 2020.

[2] E. Mayr, *Animal Species and Evolution*. 1963. Accessed: Jul. 24, 2022. [Online]. Available: https://doi.org/10.4159/harvard.9780674865327

[3] C. J. Wild and M. Pfannkuch, “Statistical Thinking in Empirical Enquiry,” *Int. Stat. Rev.*, vol. 67, no. 3, pp. 223–248, Dec. 1999, doi: 10.1111/j.1751-5823.1999.tb00442.x.

[4] Z. del Rosario, “Neglected, Acknowledged, or Targeted: A Conceptual Framing of Variability, Data Analysis, and Domain Consequences,” *J. Stat. Data Sci. Educ.*, 2024, doi: 10.1080/26939169.2024.2308119.

[5] G. Daniels, “The ‘Average Man’?,” Air Force Aerospace Medical Research Lab, Wright-Patterson AFB OH, AD010203, 1952.

[6] T. Rose, *The End of Average: How We Succeed in a World That Values Sameness*, First Edition. New York: HarperOne, 2015.

[7] L. Schiebinger *et al.*, “Inclusive Crash Test Dummies: Rethinking Standards and Reference Models,” Stanford University. Accessed: Feb. 26, 2023. [Online]. Available: https://genderedinnovations.stanford.edu/case-studies/crash.html#tabs-2

[8] GAO, “Vehicle Safety: DOT Should Take Additional Actions to Improve the Information Obtained from Crash Test Dummies,” U.S. Government Accountability Office, GAO-23-105595, Mar. 2023. Accessed: Mar. 13, 2023. [Online]. Available: https://www.gao.gov/products/gao-23-105595

[9] D. Bose, M. Segui-Gomez, ScD, and J. R. Crandall, “Vulnerability of Female Drivers Involved in Motor Vehicle Crashes: An Analysis of US Population at Risk,” *Am. J. Public Health*, vol. 101, no. 12, pp. 2368–2373, Dec. 2011, doi: 10.2105/AJPH.2011.300275.

[10] K. Hadley and W. Oyetunji, “Extending the Theoretical Framework of Numeracy to Engineers,” *J. Eng. Educ.*, vol. 111, no. 2, pp. 376–399, Apr. 2022, doi: 10.1002/jee.20453.

[11] K. Vo, A. Evans, S. Madan, and Z. del Rosario, “A Scoping Review of Engineering Textbooks to Quantify the Teaching of Uncertainty,” *ASEE Annu. Conf. Expo.*, 2023.

[12] A. J. Magana, “The role of frameworks in engineering education research,” *J. Eng. Educ.*, vol. 111, no. 1, pp. 9–13, Jan. 2022, doi: 10.1002/jee.20443.

[13] C. McCall and C. Edwards, “New Perspectives for Implementing Grounded Theory,” *Stud. Eng. Educ.*, vol. 1, no. 2, p. 93, Feb. 2021, doi: 10.21061/see.49.

[14] E. Wenger, *Communities of practice: learning, meaning and identity*. in Learning in doing. Cambridge New York Port Melbourne [etc.]: Cambridge University Press, 1998.

[15] J. L. Huff, J. Walther, N. W. Sochacka, M. B. Sharbine, and H. Kamanda, “Coupling Methodological Commitments to Make Sense of Socio-Psychological Experience,” *Stud. Eng. Educ.*, vol. 1, no. 2, p. 1, Dec. 2020, doi: 10.21061/see.29.

[16] N. W. Sochacka, J. Walther, and A. L. Pawley, “Ethical Validation: Reframing Research Ethics in Engineering Education Research To Improve Research Quality,” *J. Eng. Educ.*, vol. 107, no. 3, pp. 362–379, Jul. 2018, doi: 10.1002/jee.20222.

[17] J. Walther, N. W. Sochacka, and N. N. Kellam, “Quality in Interpretive Engineering Education Research: Reflections on an Example Study,” *J. Eng. Educ.*, vol. 102, no. 4, Art. no. 4, 2013, doi: 10.1002/jee.20029.

[18] C. Ellis, T. E. Adams, and A. P. Bochner, “Autoethnography: An Overview,” *Forum Qual. Sozialforschung Forum Qual. Soc. Res.*, vol. 12, no. 1, Art. no. 1, Nov. 2010, doi: 10.17169/fqs-12.1.1589.

[19] F. W. Ngunjiri, K.-A. Hernandez, and H. Chang, “Living Autoethnography: Connecting Life and Research,” *J. Res. Pract.*, vol. 6, no. 1, 2010.

[20] K. Charmaz, *Constructing Grounded Theory*. 2014. Accessed: Jul. 20, 2022. [Online]. Available:

https://public.ebookcentral.proquest.com/choice/publicfullrecord.aspx?p=5439903
[21] C. Faber, C. Bodnar, A. Strong, W. Lee, E. McCave, and C. Smith, “Narrating the Experiences of First-year Faculty in the Engineering Education Research Community: Developing a Qualitative, Collaborative Research Methodology,” in *2016 ASEE Annual Conference & Exposition Proceedings*, New Orleans, Louisiana: ASEE Conferences, Jun. 2016, p. 25771. doi: 10.18260/p.25771.
[22] J. C. Flanagan, “The Critical Incident Technique,” *Psychol. Bull.*, vol. 51, no. 4, 1954.
[23] J. Saldaña, *The Coding Manual for Qualitative Researchers*, 2nd ed. Los Angeles: SAGE, 2013.
[24] S. Chowdhary, S. Daitzman, R. Eisenbud, E. Pan, and E. Graeff, “Care and Liberation in Creating a Student-Led Public Interest Technology Clinic,” in *2020 IEEE International Symposium on Technology and Society (ISTAS)*, Tempe, AZ, USA: IEEE, Nov. 2020, pp. 164–175. doi: 10.1109/ISTAS50296.2020.9462188.
[25] C. Ellis, “Telling Secrets, Revealing Lives: Relational Ethics in Research With Intimate Others,” *Qual. Inq.*, 2007, doi: 10.1177/1077800406294947.
[26] M. Brewer, N. Sochacka, and J. Walther, “Into the Pipeline: A Freshman Student’s Experiences of Stories Told About Engineering,” in *2015 ASEE Annual Conference and Exposition Proceedings*, Seattle, Washington: ASEE Conferences, Jun. 2015, p. 26.1018.1-26.1018.19. doi: 10.18260/p.24355.
[27] D. Jonassen, J. Strobel, and C. B. Lee, “Everyday Problem Solving in Engineering: Lessons for Engineering Educators,” *J. Eng. Educ.*, vol. 95, no. 2, pp. 139–151, Apr. 2006, doi: 10.1002/j.2168-9830.2006.tb00885.x.
[28] R. V. Hogg, “Statistical Education for Engineers: An Initial Task Force Report,” *Am. Stat.*, vol. 39, no. 3, pp. 168–175, Aug. 1985, doi: 10.1080/00031305.1985.10479423.
[29] T. J. Chaney and C. E. Molnar, “Anomalous Behavior of Synchronizer and Arbiter Circuits,” *IEEE Trans. Comput.*, vol. C–22, no. 4, pp. 421–422, Apr. 1973, doi: 10.1109/T-C.1973.223730.
[30] R. Ginosar, “Fourteen ways to fool your synchronizer,” in *Ninth International Symposium on Asynchronous Circuits and Systems, 2003. Proceedings.*, Vancouver, BC, Canada: IEEE Comput. Soc, 2003, pp. 89–96. doi: 10.1109/ASYNC.2003.1199169.
[31] C. Wild, J. Utts, and N. Horton, “What Is Statistics?,” in *International Handbook of Research in Statistics Education*, 1st ed. 2018., D. Ben-Zvi, J. Garfield, and K. Makar, Eds., in Springer International Handbooks of Education. , Cham: Springer International Publishing : Imprint: Springer, 2018, pp. 5–36. doi: 10.1007/978-3-319-66195-7.
[32] K. Crenshaw, “Mapping the Margins: Intersectionality, Identity Politics, and Violence against Women of Color,” *Stanford Law Rev.*, vol. 43, no. 6, p. 1241, Jul. 1991, doi: 10.2307/1229039.
[33] R. V. Hogg, “Statistical Education: Improvements are Badly Needed,” *Am. Stat.*, vol. 45, no. 4, pp. 342–343, Nov. 1991, doi: 10.1080/00031305.1991.10475832.
[34] G. Box, “Statistics for discovery,” *J. Appl. Stat.*, vol. 28, no. 3–4, pp. 285–299, Mar. 2001, doi: 10.1080/02664760120034036.

Appendices

The following appendices define our reified process for making (A) and handling (B) the data through the course of this study. The initials below track which team member suggested the relevant procedure; this serves as evidence of the collaborative nature of our process definition. AG: Alex, LB: Leslie, TL: Trinity, ZDR: Zach.

A. Reflections Methods

- (LB) Read each other's reflections, produce a synthesis matrix to compare
- (AG) Establish stricter structure that is student led/driven with the ultimate goal of phasing ZDR out; focus on student reflection and mitigate power dynamics
    - Desired ZDR Role: Mediator role to keep conversation productive but doesn't dictate the content of that conversation/reflection
- (AG+LB+TL) Students meet on their own before meeting as a group with ZDR to discuss preliminary thoughts, if there are any edits they want to make to their narratives, etc.
    - Similar to PInT paper [24]

B. Reporting: Principles and Rules

Principle: Every member of the team should feel safe bringing their full, authentic self into the reflection process. No member of the team should experience a risk of personal exposure without their expressed consent.

- **(ZDR) Rule (No Verbal Discussion):** We will never discuss (verbally) the content of any other group member's reflections with a person outside this group (unless we legally have to).
- Note that this is a normal part of human subjects work, but we make this practice explicit by having this rule.

- **(ZDR) Rule (Unilateral Veto):** Every member of the team has the unilateral right to veto the disclosure of any portion of their own personal reflection to an external audience (e.g., in a publication). This is to ensure the rights and privacy of all members of the team.
- **Example:** Alice (a hypothetical team member) shared a vulnerable anecdote in-confidence, but she does not feel comfortable with that anecdote being shared outside the group. She chooses to veto including that anecdote in any external communications (including publications). The team may continue to discuss the anecdote internally, and such discussions may lead to other developments in our reflections, but the anecdote itself may not leave the group.

- **(ZDR) Rule (Requirement to Review):** Every member of the team must review any external communication of any contents of our reflections before it can be released outside the group. Combined with the unilateral veto power above, this means every member of the team has the right to remove any portion of their reflection from external communications, and it implies that every member of the team can ensure their reflections are portrayed in an acceptable way (or that individual can veto its inclusion). This review will constitute the "expression of consent" of the governing principle.

- **Example:** Bob (a hypothetical team member) shared a vulnerable anecdote, but feels that it is important to report out. Christine is working on a report of their findings, and sends a draft to Bob for review. Bob feels that Christine's portrayal of the anecdote misses some nuance, so he reaches out to talk about it. Together, they find a way to better portray the anecdote so that Bob feels the nuance is accurately communicated.

- **(TL) Rule (Content Blacklist):** Every member of the team is entitled to having a personal blacklist which outlines topics/events that are not to be shared outside of the group, unless the member has given explicit consent. This list does not replace the case by case basis of the veto or review, but it helps to clearly establish the boundaries of each team member and what they are generally not comfortable with sharing to the public. This list can also be used as a written paper trail as well as a reference point to help remind other members of what they are not comfortable with sharing.
- **Example:** Mary and John (hypothetical team members) are working on a report. John is writing up a draft which they and Mary will look over together later. Because John knows that Mary would like to avoid using her anecdotes that involve a specific topic—like her sexual orientation—through the blacklist, John knows that he should avoid including any of her anecdotes that mention that. By clearly establishing boundaries, it helps narrow down what anecdotes they can and can't use for public sharing and saves time. If John feels that there is a particularly important anecdote that falls under the blacklist, they can reach out to Mary and see if Mary would be willing to consent to that anecdote being shared. If Mary feels that they do not want it published, then they can also use the blacklist as a reference point.

Principle: All members of the team should work to protect each other's privacy and well-being.

- **(ZDR) Rule (Review for Exposure):** While conducting a review of an external communication (according to the Right to Review), all members of the team must document any information in the external communication that may constitute a risk to any member of the team. This will likely lead to multiple identifications of the same risks, but will help to ensure that no risks to any team members are missed. The entire team will review these risks together to ensure that no risks from information release are incurred without their original author's consent.