Upon completion of the interviews, a transcription was made and used to make preliminary inferences about student goals, beliefs, and conceptions of mathematical proof. The inferences were used to create narrative accounts for each student. Formal coding of the data will occur at the conclusion of the ITP course when most of the data for the larger project has been gathered.

Results

Conceptions of Proof

Most of the students were able to verbalize a description of mathematical proof compatible with proof descriptions frequently used in ITP courses. Their descriptions generally contained two parts: (1) a proof should show a statement is true or infallible and (2) a proof is made by piecing together correct mathematical statements to reach the conclusion. When asked to give a proof for number theory statements, some students were able to give arguments that were closely aligned with what may be acceptable in their course. For example, consider Max's written argument that the product of two perfect square integers is a perfect square.

"Given two perfect square integers m and n.
Let $m = k^2$ and let $n = c^2$.
Then $m \times n = k^2 \cdot c^2 = (kc)^2$.
Let $(kc)^2 = I$, integer. "

Note that Max has some issues with introducing variables (i.e. he does not give descriptors for k and l). Max felt the argument produced here was more akin to proof scratchwork than a proof. One thing he said was missing in his argument is stating the definition of perfect square. Max felt that the use of definitions was critical to a proof and even more preferrable if the definition could be given in symbols that could be manipulated. Other students had a hard time constructing an argument that resembled a proof but were able to recognize that what their arguments were far from being considered mathematical proof.

Two of the three statements given to the students were false statements. One statement was false for the entire domain of interest ("The product of two primes numbers is prime.") and one was false for part of the domain of interest ("A multiple of 3 is a multiple of 6."). With the exception of Luke, all students felt confident giving a singular counterexample to prove the falsity of a statement. Chase rewrote the false statements with explicit quantification (e.g. "All multiples of 3 are multiples of 6.") and then showed an example that showed the statement not true for all the domain. He said he was inclined to make the quantification explicit because he had heard statements of the form "for all, there is something." Rewording the statement helped Luke justify why a counterexample is sufficient evidence to prove a statement false.

Goal Orientations

While most of the students expressed a desire to earn a particular grade or pass the course, they each indicated their primary goal was the understand the complexities of proof at a deeper level by the end of the course. Max showed the most consistent evidence of have learning goal orientations. The first goal Max verbalized was to earn a high A in the course, but he rarely mentioned grades or course outcomes based on performance afterwards. He showed his eagerness to understand at a deeper level when he stated, "I just guess like a feeling of like satisfaction when it like makes sense, finally." He shared his plans to achieve his goal, including reading his textbook, attending office hours, and engaging in homework assignments and other practice opportunities. Many of these plans were based off practices he had established for success in previous mathematics class. He felt he would likely not be able to watch YouTube videos as he had in the past because he felt there would be fewer videos available to support his learning.

Luke changed his goal orientation somewhat consciously based on his interest in the material and how important he esteemed the material to be. He expressed that he began the semester focused on getting a good grade, which is indicative of a performance goal orientation. After a few days in the class he seemed to be shifting to a learning goal orientation. He said, "I think approaching it initially it was more so just get an A because it's not relevant to my major, but I find it really interesting so far. So I think my my opinion on that's going to change, I'd rather understand the subject at the—at its level." In another part of the interview, Luke said that "however it goes, it goes," which, based on the surrounding text, we infer to mean that he is not concerned with achieving a certain letter grade in the course overall. This perceived shift in goal orientation is also supported by his statement that he hopes to do better communicating with his instructor via in-person interactions to deepen his understanding, something he has not prioritized in the past.

Discussion

After less than two weeks of instruction, students' proof construction abilities were varied. Max was fairly adept constructing proofs after only 2 class sessions while Luke was not confident producing an argument resembling a proof after 3 class sessions. Since this class serves as an introduction to proof, it may be tempting for instructors to assume that all students are starting at the same base level. The contrast among student abilities at the beginning of the class is evidence that students begin the course with very different skill sets. We suggest that these differences in skill sets can be attributed, at least in part, to students' differing spontaneous conceptions of proof. Students may use their spontaneous conceptions of proof as the foundation upon which to build a scientific conception of proof. Thus, if the spontaneous foundation differs among students, the resulting scientific conceptions and skill sets will differ.

At the time of the interview, students had seen one formal proof in class (a proof that the sum of two even integers is even). The rest of the class sessions were spent discussing logical principles and the nuances of using formal logic. For example, students spoke of the difference between their spontaneous understanding of the word "or" and the new, scientific meaning of the same word. These discussions spawned in students a belief that when writing proofs, careful attention must be given to grammar, definitions, and language. Some students became so wrapped up in the scientific conception of proof that they ignored their spontaneous conceptions. For example, when deciding if an argument was a proof, they focused on grammar, language, and introduction of variables rather than the convincing or explaining component of proof.

Luke's encounters with the material during the first week of class changed his goals for the course. He began the course with a desire to earn and A and after three class sessions he was prioritizing learning the material over earning a good grade. He attributes this shift to the material being interesting. Literature on goal orientation tells us that a learning goal orientation is preferrable (Dweck & Leggett, 1988). We note that students' goal orientations may be malleable at the beginning of a course. Future research could investigate what activities led to Luke's shift to a more productive goal orientation.

The results presented here provide insights into students' initial goals, beliefs, and conceptions of mathematical proof. As we continue to interview these same seven students throughout their ITP course we will monitor how their goals and beliefs evolve from this initial starting point. This data provides a snapshot of the students at the beginning of the semester, which will compare to other data gathered throughout the semester.

References

Aberdein, A. (2009). Mathematics and argumentation. *Foundations of Science*, *14*(1), 1–8. https://doi.org/10.1007/s10699-008-9158-3

Bleiler-Baxter, S. K., & Pair, J. D. (2017). Engaging students in roles of proof. *The Journal of Mathematical Behavior*, *47*, 16–34. https://doi.org/10.1016/j.jmathb.2017.05.005

Czocher, J. A., & Weber, K. (2020). Proof as a cluster category. *Journal for Research in Mathematics Education*, *51*(1), 50–74. https://doi.org/10.5951/jresematheduc.2019.0007

David, E. J., & Zazkis, D. (2020). Characterizing introduction to proof courses: A survey of U.S. R1 and R2 course syllabi. *International Journal of Mathematical Education in Science and Technology*, *51*(3), 388–404. https://doi.org/10.1080/0020739X.2019.1574362

Dawkins, P. C., & Weber, K. (2017). Values and norms of proof for mathematicians and students. *Educational Studies in Mathematics*, *95*(2), 123–142. https://doi.org/10.1007/s10649-016-9740-5

Dweck, C. S., & Leggett, E. L. (1988). A social-cognitive approach to motivation and personality. *Psychological Review*, *95*(2), 256–273.

Hanna, G. (2000). Proof, explanation and exploration: An overview. *Educational Studies in Mathematics*, *44*(1), 5–23. https://doi.org/10.1023/A:1012737223465

Herbst, P. G. (2002). Establishing a custom of proving in American school geometry: Evolution of the two-column proof in the early twentieth century. *Educational Studies in Mathematics*, *49*(3), 283–312. https://doi.org/10.1023/A:1020264906740

Hoffman, K., Williams, T. H., & Kephart, K. (2024). The use of guided reflections in learning proof writing. *Education Sciences*, *14*(10), Article 10. https://doi.org/10.3390/educsci14101084

Moore, R. C. (1994). Making the transition to formal proof. *Educational Studies in Mathematics*, *27*(3), 249–266. https://doi.org/10.1007/BF01273731

Morris, A. K. (2007). Factors affecting pre-service teachers' evaluations of the validity of students' mathematical arguments in classroom contexts. *Cognition and Instruction*, *25*(4), 479–522. https://doi.org/10.1080/07370000701632405

Stylianides, G. J., Stylianides, A. J., & Weber, K. (2017). Research on the teaching and learning of proof: Taking stock and moving forward. In *Compendium for research in mathematics education* (pp. 237–266).

Vygotsky, L. S. (1987). *The collected works of LS Vygotsky: The fundamentals of defectology* (Vol. 2). Springer Science & Business Media.

Values in Tension: Mathematics Teacher Educator's Decisions on Generative AI

Sarah D. Castle
University of Idaho

Marcus Wolfe
James Madison University

Hyunkyoung Yoon
California State Polytechnic University, Pomona

Shahabeddin Abbaspour Tazehkand
Missouri State University

This analytic autoethnography preliminary report examines how early-career mathematics teacher educators (MTEs) make decisions about if and how to use generative AI (GenAI) in alignment with their values. Using reflective journals and group discussions, with Value-Sensitive Design, our analysis highlights three themes in GenAI use: preserving autonomy and creativity, protecting human elements of our job while viewing GenAI as a supplemental tool, and managing institutional pressures. This work highlights values-based perspectives informing the broader understanding of MTE GenAI decision making.

Keywords: GenAI, Values, Professional Development, Ownership, MTEs

Emerging research focuses on how generative AI (GenAI) impacts student thinking and learning, ethical considerations surrounding the infrastructure needed for large language models (LLMs), and classroom practices. Critics in education have raised concerns over GenAI's impact on students' perception on mathematics, critical thinking skills, autonomy, and problem solving (Kosmyna, 2025; Lee et al., 2025; Schorcht et al., 2024; Wang et al., 2024). Further, researchers found that students need support in determining how to critically engage with these tools and not passively accept their outputs (Yoon et al., 2024), and teacher educators specifically support preservice teachers' developing critical dispositions (Sawyer & Aga, 2024). However, far less is known about how teacher educators navigate GenAI, especially for those early in their career who are still finding their professional identity. Their choices not only inform their practice but set norms for preservice teachers about how to integrate, or resist, GenAI in their classrooms.

Understanding how teacher educators make choices about GenAI humanizes the broader conversation about its integration in education. During this period of sociocultural shifts in response to GenAI, we as educators should be transparent about when and if we use GenAI. This is particularly important for mathematics teacher educators (MTEs) as mathematics emphasizes reasoning, justification, and sensemaking which can be undermined if GenAI produces answers without engaging in mathematical practices. Given the lack of professional guidelines, early career MTEs are facing nebulous choices about GenAI in their teaching and professional development while navigating their induction into the profession. As a group of young MTEs, we ask the question: *How do we as mathematics teacher educators make decisions to align with our basic values about using GenAI in our teaching, scholarship, and service?* To answer this research question, we use analytic autoethnography methodology coupled with Value-Sensitive Design (VSD) to understand the different tensions that emerge when early career MTEs navigate GenAI use. By examining our own experiences and reflections, this preliminary study aims to uncover the core values MTEs draw on when making choices around the use of GenAI.

Positionality Statement

We are four early career mathematics teacher educators who share a commitment to preparing ambitious mathematics teachers, supporting in-service teachers, and contributing to research-based interventions. We identify across a range of racial and gendered identities as we are composed of one Middle Eastern man, one Asian woman, one Black man, and one White woman. Professionally, we represent R1, R2, and Research Colleges and Universities. We are located across the US in education and mathematics departments. Collectively, we bring our different lived experiences to the research as we have nuanced dialogue that explores the similarities and differences in our approaches to GenAI in mathematics teacher education.

Theoretical Framing

We draw on Value-Sensitive Design (VSD), a theory and methodology used to provide a systematic process to account for human values in technology development within a multi-stakeholder system (Friendman et al., 2002). VSD positions technology as non-neutral since developed technology will always be intertwined with human values. These values manifest in layers, and within education technology we can understand values as basic values, learning values, or technology use values (Prieto et al., 2025) which should be iteratively coded. This allows us a lens for understanding the interaction between GenAI and our values, as well as to look across multiple layers of stakeholders in education, both direct and indirect (Shen et al., 2025) because as MTEs, we are in within a department that serves students within a university within academia. We can trace both how and why we make decisions about using GenAI by uncovering the personal and professional values that guide those choices. These choices are not simply pragmatic but are intertwined with our professional identities, institutional pressures, and underlying commitments as MTEs. For value categorization we used Schwartz's (2006) core-values. Under this theory of basic values, values are desirable goals that motivate action that often serve as standards or criteria. People's values dictate an ordered system of priorities and the relative importance of multiple values guide action. There are ten core values given: self-direction, stimulation, hedonism, achievement, power, security, conformity, tradition, benevolence, and universalism.

Methodology

We used analytic ethnography to investigate the personal and professional values that guide how we, as MTEs, chose to use (or not use) GenAI. This type of ethnography is where the researcher is "(1) a full member in the research group or setting, (2) visible as such a member in published texts, and (3) committed to developing theoretical understandings of broader social phenomena" (Anderson, 2006). This methodology focused on our roles as both researcher and participant, thereby allowing for reflexivity in the analysis process and situating our narratives in a broader theoretical and institutional context. Data collection consisted of iterative individual reflective journals as well as group discussions over a three-month period. The journal prompts focused on our feelings and choices surrounding GenAI in our own teaching, service, and research. After the initial journaling, we read each other's narratives and rejoined for a group discussion where we could ask follow-up questions and focus on what was shared among members and where our approaches diverged. All meetings were transcribed and then we engaged in an iterative coding process to engage in individual sense-making and collective meaning-making. We finally coded all data against Schwartz (2006) core-values for our preliminary VSD basic values and proceeded to group codes and look for themes.

Preliminary Results

Our analysis revealed three key themes, each reflecting core values, namely self-direction, security, and ownership, that shaped our decision about GenAI use: (1) Preserving autonomy and creativity when deciding whether to use GenAI, (2) recognizing that GenAI cannot replace the fundamentally human parts of being an MTE, but could serve as a supplemental tool, and (3) GenAI is a way to negotiate institutional demands that we do not view at the core of teaching.

Self-Direction

A core value that we drew on when deciding whether or not to use GenAI was self-direction, which focuses on independent thought and action, as well as choosing, creating, and exploring. This value manifested within our data as our commitments to autonomy, critical thinking, and creativity in our teaching and research. Within the classroom, we positioned GenAI as a tool for exploration and problem-solving if used, rather than shortcutting mathematical reasoning. We chose to allow for the use of GenAI in this manner because we wanted to have students make choices about their learning and have the opportunity to critically engage with GenAI and discourage simply copying and pasting answers. This reflected our commitment to autonomy as students are encouraged to direct their own learning and think both critically and independently. Within our research, we chose not to have GenAI independently generate references or citations for our work as this was in violation of autonomy, and we value self-direction. By choosing our resources or engaging in a systematic review process that we designed, we maintained our self-directed inquiry and focused on how we actively shaped our research rather than outsourcing intellectual labor to a machine. This focus on research and intellectual labor centered a larger conversation on the humanization of research and what does knowledge and ownership look like when human and machine contributions overlap?

Ownership and ideas. Ownership emerged as a central theme and value that shaped how we chose to use, or not use, GenAI. As ownership of ideas focused on making decisions about what to create, explore, or think critically about, we viewed it as a basic value under self-direction as it fundamentally centers on our autonomy over thinking and contributions. There was concern about not only what was produced but who was responsible for creating knowledge. This value informed decisions across multiple aspects of our work. In teaching, many of us described clear boundaries around students' work. Entering student assignments and work into AI tools felt like a breach of ownership and trust, since students shared their work with us, but never consented to AI, and this work belongs to the student. This demonstrates that ownership is not simply an abstract principle, but a value that is operationalized in a lived ethical boundary embedded in daily pedagogical practice.

Within our professional writing and communication, the value of ownership was often in tension with the need for efficiency within the academic system. Many expressed concern that over relying on GenAI would weaken their professional voice, thereby hiding identity and authorship from their writing. This led to us heavily editing AI outputs when used, and to treat them as starting points, rather than the final product. This allowed us to make decisions aligned with our desire for self-direction while maintaining productivity. However, many expressed the concern on homogenization of ideas and writing tone when using GenAI and wanting to promote individual thought and diverse ideas and writing styles.

One of the clearest choices made around GenAI as MTEs was protecting the humanization of ideas and roles. We consistently drew lines around activities that we considered core to our professional role and identity. Teaching and research, in particular, were framed as fundamentally human endeavors. As one reflection noted, education must remain a “human endeavor” where feedback, instruction, and inquiry flow from authentic human engagement, not automated outputs. This manifested within research as many of us expressed reluctance to use AI tools within qualitative analyses, as having a weighted algorithm categorizing, coding, and grouping thereby undermines the reflexive research process and may cause harm. However, for elements of research that were not necessarily viewed as essential or relational, like debugging code, checking grammar, or organizing references, we were much more inclined to use AI tools, so long as the central intellectual labor was still-human driven. These instances highlighted how our effort to preserve ownership and autonomy in thinking forces GenAI tools to be viewed as a supportive tool rather than replacement of work.

Our choice of whether to use GenAI tools was not solely confined to our own work, but also for students. We raised concerns that if students turned to GenAI too quickly for answers, they would "cheat themselves" out of the process of thinking. It was important to all of us to emphasize the importance of cultivating critical evaluation skills and perseverance, so students could use GenAI as a dialogue partner rather than a replacement for their own reasoning. Allowing students to co-construct classroom AI policies was seen as one way of reinforcing ownership, making explicit the terms under which GenAI would or would not be used.

Ultimately many of us discussed feelings of guilt surrounding the use of GenAI when tasks were within our professional capacity because there was a feeling of inauthenticity as well as lack of ownership. Further, we were deeply troubled with current research highlighting the psychological impact of ownership and GenAI (Kosmyna, 2025; Lee et al., 2025). The more secure individuals felt in their intellectual capacities, the less likely they were to outsource critical thinking to machines. On the other hand, lack of confidence could lead to over-reliance, undermining ownership of ideas and even memory of the work produced. We had greater concerns of our GenAI use as early career MTEs, as this stage of academic work is often associated with imposter syndrome and developing our own professional identity.

Conformity and Security

Our values were not in isolation, rather our values of conformity and security interacted with the universities' institutional values, such as achievement and power. As early career MTEs thinking about their professional identity, we often focused on our belonging within professional communities and stability within our job and research; however, this was in competition with the expectations that are placed upon 'measurable' achievements and productivity. Within our context, we operationalized conformity as making decisions to not violate expectations of us as MTEs. These social expectations came from our universities, professional communities, and students. As mentioned, this was often tied with security, which focused on our safety within our job and the research community. Simultaneously, the university values of power and achievement impacted our decisions around GenAI as universities wanted to maintain a rigorous professional image, in alignment with power, and many times used physical measures that are social standards to determine achievement.

Due to the focus on measurable outcomes and the social norms imparted on us, and out of a sense of wanting stability in our profession, this created a context in which GenAI was a pragmatic strategy that was used to reconcile competing value demands. Often GenAI is what allowed us to meet the needs of the university and conform to expectations while being able to preserve what we viewed as the heart of teaching. When our commitment to our students, our research, or other job elements did not align with the universities' measures of achievement, this led to a greater likelihood of using AI. AI was not simply a technological tool but rather was a way to try and protect our core teaching and research values while meeting social norms and keeping our jobs. Even if using AI was not consistent with personal values like self-direction or universalism, including environmental concerns, some resulted to AI to meet these needs. For example, consider one of our experiences navigating AI use:

> "I struggle because it feels like I should be able to do all these things on my own and yet trying to do it all is not physically possible… I am trying to prioritize students… Many times, for generic things that I have to do as a faculty member that no one really reads, I will have ChatGPT do a first draft because if no one else is really caring about it, then why should I put my time and effort into it, especially if it is basically adult busywork. However, why should I make my life easier at the expense of others?"

Within this quote, it is evident that this MTE wants to focus on students; however, the university values of achievement and their desire for conformity created a situation where they felt they had to use GenAI. This choice was inconsistent with their value of universalism, but the value of security outweighed this concern.

The time pressure related to academic work and the change in communication expectations were expressed across participants and in these situations, our values combined with university values, created a space where GenAI allowed us to maintain our commitment to teaching while meeting the social expectations. A common situation was with regards to email and electronic communications. This practice was important to the university, but we did not feel was at the core of teaching. One member noted that "the need to communicate consistently, and keep the right tone, and the fact that I am not confident in this skill is the reason I have used AI to help me rephrase my email drafts." Another stated that "one of the things that I feel like that has changed, especially since COVID is the expectation of speed and time to reply to things. And so the only way for me sometimes to meet those deadlines with emails [which] to me it isn't at the core of teaching [is to use AI]". These are specifically tied to being "[measured] in number of articles or something like that or number of emails sent" which are norms set by the universities and professional communities. Further, despite the desire of universities to have achievement and power, one member, with agreement of others, expressed frustration in the university not willing to provide resources. Our reflections highlight that the use of GenAI in our work is more of a values-mediated coping tool as we respond to additional institutional and societal pressures related to tasks we do not highly value or view as innately human. It is not that we wanted to actively ignore GenAI, rather, to maintain the critical thinking desired in the classroom and to have time for mentoring students and conducting research, there are areas of our job that we were willing to compromise and use GenAI to complete. These areas are what the institution may deem necessary but are not at the core of our professional values.

Moving Forward

This preliminary report focuses on the first layer of basic value coding with our initial data. While we have identified some personal values and how they interact with institutional values over time, the application of VSD is still being worked on, particularly when we think of the interaction and creating a framing of the implications on GenAI use across all values. Further, we view this preliminary report and subsequent discussion with the larger community and stakeholders as part of the analytic autoethnography process. Questions that we plan to pose to the audience include (1) How could VSD be adapted to account for evolving professional expectations when core values are fixed? (2) Are there different values or tensions in the VSD framework that attendees feel we overlooked either in our analysis or based on their experience? (3) How can this analytic autoethnography serve a broader MTE audience for GenAI decision making? Feedback from attendees will further refine our framing and analysis while ensuring that this autoethnography is linking our experiences to the broader phenomena of MTE GenAI decision making. This feedback will help us explore the subsequent layers of VSD with learning and technology values.

References

Friedman, B., Kahn, P., & Borning, A. (2002). Value sensitive design: Theory and methods. *University of Washington technical report*, 2(8), 1-8.

Friedman, B., Kahn, P., Jr., & Borning, A. (2006). Value sensitive design and information systems. In P. Zhang & D. Galletta (Eds.), From *human-computer interaction and management information systems: Foundations* (Advances in Management Information Systems, Vol. 5, pp. 348–372). M.E. Sharpe. https://doi.org/10.1007/978-94-007-7844-3_4

Kosmyna, N., Hauptmann, E., Tong Yuan, Y., Situ, J., Liao, X., Vivian Beresnitzky, A., Braunstein, I., & Maes, P. (2025). Your brain on chatgpt: Accumulation of cognitive debt when using an ai assistant for essay writing task. arXiv preprint arXiv:2506.08872.

Lee, H.-P., Sarkar, A., Tankelevitch, L., Drosos, I., Rintel, S., Banks, R., & Wilson, N. (2025, April 26–May 1). The impact of generative AI on critical thinking: Self-reported reductions in cognitive effort and confidence effects from a survey of knowledge workers. In *Proceedings of the CHI Conference on Human Factors in Computing Systems* (CHI '25) (Yokohama, Japan). Association for Computing Machinery (ACM). https://doi.org/10.1145/3706598.3708414

Prieto, L. P., Viberg, O., Yip, J. C., & Topali, P. (2025). Aligning human values and educational technologies with value-sensitive design. *British Journal of Educational Technology*, 56(4), 1299–1310. https://doi.org/10.1111/bjet.13602

Sawyer, A. J., & Aga, Z. (2024). Counterexamples to demonstrate artificial intelligence chatbot's lack of knowledge in the mathematics education classroom. *Connections: The Newsletter of the Association of Mathematics Teacher Educators*, 24(1), 1–4.

Schorcht, S., Buchholtz, N., & Baumanns, L. (2024). Prompt the problem – Investigating the mathematics educational quality of AI-supported problem solving by comparing prompt techniques. *Frontiers in Education*, 9, 1386075. https://doi.org/10.3389/feduc.2024.1386075

Shen, Y., Tang, L., Le, H., Tan, S., Zhao, Y., Shen, K., Li, X., Juelich, T., Wang, Q., Gašević, D., & Fan, Y. (2025). Aligning and comparing values of ChatGPT and human as learning facilitators: A value-sensitive design approach. *British Journal of Educational Technology*, 56(4), 1391–1414. https://doi.org/10.1111/bjet.13562

Wang, K. D., Wu, Z., Tufts, L., Wieman, C., Salehi, S., & Haber, N. (2025). Scaffold or crutch? Examining college students' use and views of generative AI tools for STEM education. In *Proceedings of the 2025 IEEE Global Engineering Education Conference (EDUCON)* (pp. 1–10). IEEE. https://doi.org/10.1109/EDUCON62633.2025.11016406

Yoon, H., Hwang, J., Lee, K., Roh, K. H., & Kwon, O. N. (2024). Students' use of generative artificial intelligence for proving mathematical statements. *ZDM – Mathematics Education*, *56*(7), 1531–1551. https://doi.org/10.1007/s11858-024-01629-0