Constructive Dialogue Institute

Can AI Teach Dialogue Skills to College Students?

Constructive Dialogue Institute (CDI)

Founded in 2017, CDI is a non-profit organization dedicated to equipping the next generation of Americans with the mindset and skill set to engage in dialogue across differences. At CDI, we seek to help teachers, faculty, and administrators build learning environments that enable students to feel comfortable engaging with challenging topics so that real learning can occur. To accomplish this goal, we translate the latest behavioral science research into educational resources and teaching strategies that are evidence-based, practical, and scalable.

Authors

Ryan W. Carlson, PhD — Research Scientist

Constructive Dialogue Institute

www.constructivedialogue.org
244 Madison Ave, Suite 1098
New York, NY 10016

Publication date: March 2026

Constructive Dialogue Institute

Contents

Introduction

2026 is an inflection point for higher education. Political polarization is near historic highs as Americans grow divided not just over their values, but over basic facts.[1–3] At the same time, generative AI is reshaping how students learn, communicate, and engage with ideas as hundreds of millions of college-aged adults now use AI every week.[4] Universities are navigating both challenges, as they seek to engage with AI responsibly while equipping their students—the next generation of citizens and leaders—with the tools to bridge the divides our country is facing.

Emerging evidence suggests these two challenges might intersect in significant ways. AI tools, when thoughtfully designed, can play a meaningful role in supporting constructive dialogue across differences. But the same qualities that make AI tools powerful, such as their responsiveness and persuasiveness, also introduce risks that make careful design essential.

This report reviews the emerging evidence on how generative AI can support constructive dialogue on college campuses and beyond.

Our core recommendation for deploying AI in constructive dialogue

A 2026 survey of over 4,000 students and faculty shows 95% already use AI in educational contexts, yet only one in four institutions has a formal policy governing that use.[5] The decision for higher ed leaders is no longer whether to adopt AI on campus, but how to shape its use.

This report reviews emerging evidence on one important question within this landscape: how AI can be used to support constructive dialogue. We identify three emerging roles:

- **AI as a coach** for building individual dialogue skills
- **AI as a mediator** for facilitating conversations across differences
- **AI as a conversation partner** for engaging students in disagreement

The evidence reviewed here points to a clear pattern: the more constrained and pedagogically structured the AI tool's role is, the lower the risk and the stronger the current evidence for its use. We recommend campus leaders:

- **Start with AI coaching,** where the benefits are best-established and the risks are most manageable
- **Expand cautiously to AI mediation,** with appropriate safeguards for live conversations
- **Approach AI conversation partners with greater caution,** since this use case presents the highest level of risk

Across all three roles, purpose-built AI dialogue tools allow institutions to offer structured alternatives to generic AI, with pedagogical design, safeguards, and oversight built in.

The Problem and Promise of AI for Constructive Dialogue

Constructive dialogue is a form of conversation in which people with differing views seek to understand one another, without abandoning their own beliefs, in order to engage productively across differences.[6]

To understand the role AI can play in supporting constructive dialogue, it helps to first consider three central challenges students face in engaging in dialogue across lines of difference.

1. **Dialogue feels risky.** Communicating across differences is difficult, and this feeling is especially potent in college, an intensely social experience that makes constructive engagement feel reputationally risky. When conflict arises, it can feel easier to pick a side or to disengage entirely. The social rewards of expressing outrage often outweigh those of engaging across differences.[7] For students with minority viewpoints, the risks of speaking up can feel tangible and consequential, making silence the safest response.
2. **Practice opportunities are hard to come by.** Even if students are willing to engage constructively across differences, they may struggle to find people in daily life with whom they genuinely disagree. Students naturally cluster with like-minded peers, a phenomenon known as homophily,[8] which limits access to people who hold different views. Yet these encounters are necessary to strengthen their capacity for dialogue.
3. **Dialogue requires deliberate practice that students rarely receive.** Constructive dialogue rests on a set of concrete skills such as active listening,[9] perspective-taking,[10] asking open-ended questions,[11] and deeper capacities like intellectual humility[12] and emotional regulation.[13] These skills and capacities can be cultivated through deliberate practice.[14] Yet most students never receive structured guidance, opportunities to practice component skills, or access to coaching and feedback.

AI holds promise to address each of these constraints. Generative AI tools can create a low-stakes space for engaging across differences, removing the social risk that drives avoidance. Researchers are using AI to simulate disagreement, solving the exposure problem that homophily creates. And new tools are being developed to deliver structured, personalized practice—breaking dialogue into component skills, providing granular feedback, and adapting to where each learner is struggling. Crucially, these tools can do this at scale, reaching students who would otherwise never encounter structured dialogue practice at all.

Three Emerging Roles for AI in Dialogue

Emerging research suggests that AI can support constructive dialogue through three distinct roles: as a *coach*, *mediator*, or *conversation partner*. Each role has its unique advantages and risks, and understanding how they differ is important for deploying AI wisely.

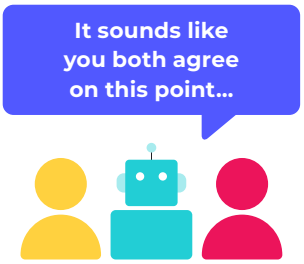
AI coaches work with students privately to build dialogue skills through targeted practice and real-time feedback. An AI coach's purpose is primarily *pedagogical*: it teaches students how dialogue works by prompting them to reflect, evaluating their technique in scripted dialogue scenarios, and withholding answers to ensure students experience the friction necessary for learning. Its role is limited to skill development within defined exercises created by researchers or educators.

AI mediators facilitate live conversations between real people, supporting the *process* of dialogue. An AI mediator's purpose is not to contribute its own viewpoints, but to structure an exchange between people by prompting elaboration, surfacing agreement, and redirecting drifts in conversation.

AI conversation partners engage directly with students, offering a low-stakes space to practice disagreement with a realistic counterpart. Unlike coaches or mediators, conversation partners generate arguments, counterarguments, and rebuttals in real time, directly engaging with a student's beliefs and positions.

Each of these roles for AI holds promise to address the three barriers described earlier: social risk, limited exposure, and the skill gap. Yet they achieve it in different ways and to different degrees. Coaching primarily targets skill development while indirectly lowering social risk through private practice. Mediation mainly reduces the interpersonal risk of a live conversation while modeling some principles of constructive dialogue indirectly. Conversation partners most directly address the exposure barrier by simulating disagreement on demand, while also creating a low-stakes environment.

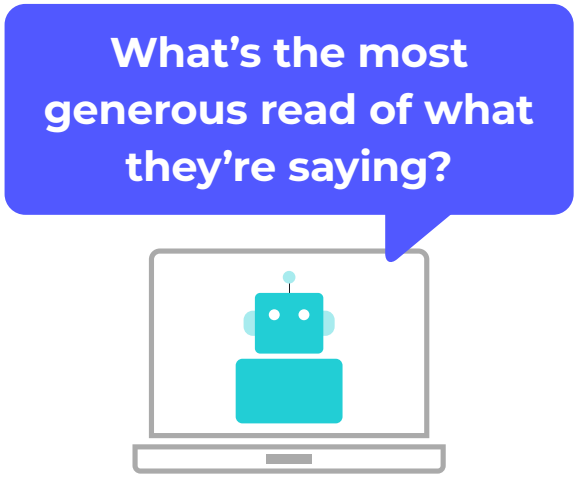
The nature of the risks shifts by role as well. For coaches, the central question is whether the tool is designed well enough to build real skills. For mediators, it's whether the AI can responsibly manage a live conversation between real people. For conversation partners, the risks grow more serious, including the capacity to shift beliefs or distort understanding of those we disagree with.

Three roles for AI in Constructive Dialogue			
Role	**AI as Coach**	**AI as Mediator**	**AI as Conversation Partner**
			
Assessment	**Most constrained role, lowest risk**	**Moderately constrained role, moderate risk**	**Least constrained role, highest risk**
Primary Function	Builds skills through structured practice and feedback	Facilitates dialogue between humans	Practice conversation partner that simulates disagreement
Key Benefits	Reduces social risk, creates practice opportunities, and develops skills through personalized feedback	Makes real conversations safer and more productive through facilitation	Creates low-stakes exposure to disagreement
Key Risks	Risk of over-scaffolding, students may develop false mastery, feedback may enforce narrow cultural norms	Students may not internalize dialogue skills, risk of false equivalence, oversight challenges when conversations escalate	Risk of persuasion, practice without feedback may not build transferable skills, risk of caricaturing opposing views, generates the most sensitive conversational data

AI as Coach: Most Constrained Role, Lowest Risk

AI coaches are like personal trainers for dialogue. They work with students privately, strengthening skills through targeted practice and real-time feedback. They reflect the most constrained role AI can play in dialogue, and also carry the lowest risks to manage.

The core promise of AI coaches is that they enable *deliberate practice*. Rather than sending students directly into difficult conversations, coaches allow students to isolate and strengthen individual skills like active listening, asking open-ended questions, and generating charitable interpretations of opposing views.



Evidence of skill acquisition. Research on AI coaching is young but growing, and initial findings are encouraging. AI coaches have helped second-language learners improve speaking proficiency,[15] mental health workers respond more empathetically,[16] and students show greater academic achievement.[17] Moreover, AI coaching with structured feedback improves interpersonal skill mastery, self-efficacy, and emotion regulation in challenging conversations.[18] Together, these findings suggest that AI-driven coaching can produce measurable gains across a range of skills and settings.

These findings align with what students themselves report wanting. Our own research suggests that students are eager to improve but typically lack structured feedback loops. Without this guidance, they rely on self-reflection or peer reactions, which often lack the clarity needed for real progress.[19]

What separates effective AI coaches from generic AI tools is deliberate pedagogical design. Bastani et al. offer a striking illustration: students using out-of-the-box GPT-4 as a math tutor showed a performance *decline* on subsequent exams, while students using a custom GPT tutor with pedagogical guardrails did not.[20] The implication: an AI coach that hasn't been intentionally designed can undermine the very learning it aims to support. The same principle applies to dialogue coaching. A coach that simply validates students' responses, or does the thinking for them, may produce the illusion of skill without the substance.

The risks of AI coaches are manageable. Unlike AI mediators and conversation partners, the risks here are primarily pedagogical. The central question is whether the tool is designed carefully enough to actually build the skills it promises, rather than substituting for them. The risks we outline below are real, but highly addressable through careful design.

Over-scaffolding. As with any form of coaching, AI coaches can provide too much structure. Too little structure, as Bastani et al. suggest, leads students to let AI do the thinking for them.[20] But too much structure carries its own risk: if a coach offers hints before students have struggled, or constrains responses too narrowly, it can also become a crutch.

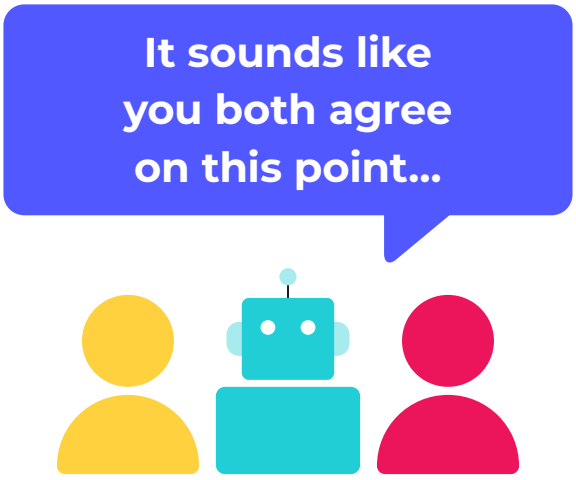
Illusions of understanding. Merely working with AI can give rise to illusions of understanding,[21] where people believe they understand more than they do about a topic. Students who receive AI feedback about dialogue may feel more confident after these interactions without actually internalizing skills. This misperception transforms into a risk when real conversations arise: a student who overestimates their ability may enter conversations less prepared and less humble than before they practiced.

Tone policing. Without cautious design, AI coaches may be too rigid in the expressions of constructive dialogue they accept and praise. Students' cultural backgrounds can give rise to diverse ways of expressing constructive dialogue principles, and an AI coach trained too narrowly may enforce stylistic norms rather than develop genuine skill. Recent research has shown that LLMs can form implicit associations that influence downstream decisions, even when the same models pass standard bias benchmarks.[22] For campuses deploying AI coaches, it's important to ensure coaches are designed with feedback rubrics that account for a range of cultural backgrounds.

Mitigating these risks is achievable through carefully designed prompts and guardrails: preserving the friction needed during practice to strengthen skills, adapting difficulty to each student's level, and building in culturally responsive feedback rubrics. These risks require attention, but they can all be addressed.

AI as Mediator: Moderately Constrained Role, Moderate Risk

Whereas AI coaches develop an individual's skills in private, a well-designed AI mediator steers the process of dialogue between people, prompting elaboration, surfacing common ground, and keeping discussions on the rails. By reducing the friction of live disagreement, AI mediators can lower the motivational barriers to engagement. AI mediators reflect a moderately constrained role for AI in dialogue and carry a moderate degree of risk to manage.



We focus on evidence from two tools as case studies: the Habermas Machine, developed by researchers at Google DeepMind,[23] and Sway, a platform developed by researchers at Carnegie Mellon University.

Large-Scale Evidence: The Habermas Machine. Can AI help people with genuinely different political views find common ground? Researchers at DeepMind gave groups of participants a divisive political topic, such as immigration or Brexit. Each participant's private view was fed to an AI mediator (the Habermas Machine), which generated a group consensus statement, collected group members' critiques, and revised the statement over multiple rounds. Across nearly 6,000 UK participants, groups preferred AI-generated consensus statements over those written by human mediators, rating them as clearer and more informative.[23]

The most striking finding came from the control condition: groups that simply exchanged opinions without AI mediation showed no change in agreement at all. By contrast, with AI mediation, the proportion of groups reaching unanimous agreement nearly doubled.

The Habermas Machine's role was restricted to synthesizing diverse views rather than live facilitation, so participants never spoke directly to one another. It thus demonstrates AI's capacity to identify shared ground, but offers less direct evidence for AI mediation of real-time conversations. This is what Sway tests.

Preliminary Evidence: Sway. How does AI mediation perform in one-on-one dialogue? Sway is a chat platform that connects participants across differing perspectives while an AI mediator supports the conversation by prompting elaboration, surfacing areas of agreement, and looking up factual information at users' request.

In an early pilot study, 120 participants were paired with someone holding opposing views on polarizing topics. After one 30-minute conversation, participants reported a greater ability to understand someone they disagreed with, and a greater openness to opposing views.[24] The results are preliminary, but they suggest that AI mediation can make even difficult one-on-one conversations more productive.

The risks of AI mediators are moderate. Unlike AI coaches, AI mediators facilitate real, dynamic disagreements between people, which raises the stakes. In addition to pedagogical concerns (for example, dependency on the mediator), this role introduces risks that touch on how AI represents reality and manages a live, shared situation. These risks include AI fact-checking errors, false equivalence in debate, and the challenge of human oversight when conversations escalate. Some of these risks are empirically documented; others are grounded in established principles of conflict mediation that are important to anticipate.

Mediator dependency. Since AI mediators are designed to facilitate rather than teach dialogue skills, they risk producing conversations that succeed without students ever acquiring the underlying dialogue skills. The mediator supports the process, but students may never internalize *why* it worked. This is a pedagogical limitation; mediation can build the conditions for people to have a productive dialogue without building peoples' capacity to replicate it independently.

AI as arbiter of truth. Some AI mediators also check factual claims in real time to help resolve disagreements. This is a useful function, but it assumes AI can reliably distinguish fact from misinformation. Because inaccuracies are embedded in AI training data, there is a potential risk that AI fact-checking on controversial topics can introduce errors, rather than correct them.[25]

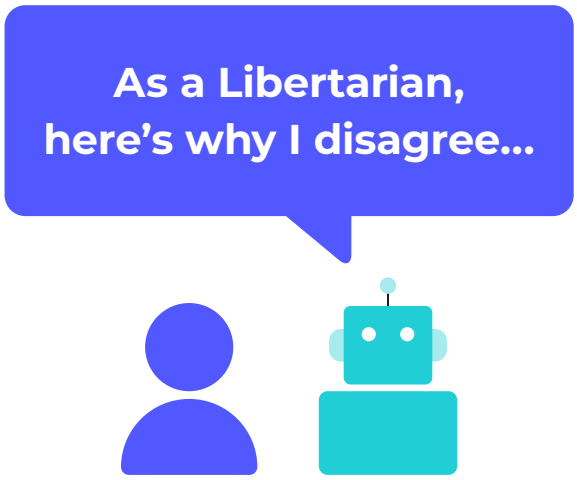
False equivalence. When an AI mediator treats all positions as an equally legitimate starting point, it risks creating false equivalence. Not all claims carry the same degree of evidence or factual grounding, and by smoothing over differences too aggressively in pursuit of consensus, an AI mediator may present an empirically unsupported view as comparable to a well-substantiated one. In educational settings, this can blur the distinction between evidence-based arguments and unsupported assertions. A trained human mediator can read the room when these situations arise, and adjust accordingly. It is not yet clear that AI mediators can make such judgments reliably.

The oversight question. Since mediation involves real people interacting in real time, the oversight concerns are qualitatively different from private coaching. A coaching session gone wrong affects one student privately; a mediated session gone wrong affects multiple students, potentially in a shared space. Ideally, live AI-mediated sessions should include a protocol for human facilitator intervention when conversations escalate, but where that isn't feasible, practitioners should weigh the risks of proceeding against the potential benefits.

These risks are real, but they're not inherent to AI mediation; they're consequences of inadequate design and guardrails. Platforms that address them thoughtfully, with safeguards against false equivalence, inaccurate fact-checking, and escalation, represent a meaningfully safer option than leaving live disagreement entirely unmediated or in the hands of generic AI tools.

AI as Conversation Partner: Least Constrained Role, Highest Risk

AI conversation partners are arguably the most intuitive of the three roles, but carry the greatest risks to mitigate. The appeal is straightforward: like a flight simulator for dialogue, they promise to offer realistic practice for high-stakes conversations without social consequences. Students can encounter pushback and counterarguments without fear of losing friends or face. Unlike coaching, where educators design structured scenarios, AI conversation partners respond dynamically and unpredictably, mirroring what real disagreement actually feels like.



That appeal is also the problem. AI conversation partners are compelling because they seem to address two barriers at once, offering a low-stakes space to build skill while removing the social risks that deflate students' motivation to engage. But the evidence of translatable skill gains from these interactions is currently lacking, and the risks are substantial.

The risks of AI conversation partners are high. AI conversation partners are the least constrained of the three roles by design. Simulating realistic disagreement requires these tools to generate arguments and counterarguments in real time, with fewer guardrails on what they can say or how persuasively they can say it. The qualities that make AI conversation partners convincing are the same ones that make them difficult to constrain. As a result, the risks here go beyond pedagogy: they include the capacity to actively shift students' beliefs, reinforce distorted views of those we disagree with, and expose sensitive personal data, among others.

The persuasion risk. In just three conversational turns, an AI conversation partner reduced conspiracy theorists' beliefs in a discussed conspiracy by 20% on average, with effects persisting two months later.[26] To some, this may seem like a positive result, but AI conversation partners can change beliefs in any direction.[27] AI outperforms human persuaders,[28] and conversational AI advocating for political candidates can significantly shift voter preferences, sometimes on the basis of inaccurate claims.[25] This persuasive capacity is core to what makes AI a compelling conversation partner: its responsiveness, its fluency with evidence, and its persistence.

Constructive dialogue is not about persuading others; it's about mutual understanding. Designing a simulated conversation partner that's realistic enough to be useful but not persuasive enough to durably shift beliefs is a tension that current tools have not resolved.

The limits of practice without feedback. Engaging in a conversation with AI does not, by itself, translate into skill development. One illustrative example comes from research on negotiation training. Researchers found that participants using the ACE system improved only when they received targeted feedback between rounds of negotiating against an AI counterpart.[29] This work mirrors decades of research showing that experience without instruction produces limited skill gains.[30] Without an accompanying coach, practice against a realistic AI counterpart may not build transferable dialogue skills.

Caricature reinforcement. AI conversation partners also risk reinforcing caricatured views of those we disagree with. An AI simulating a political opponent draws on training data shaped by media portrayal and online discourse, which tend toward the extreme and stereotypical. Without deliberate design to present the strongest version of opposing positions, students practice against a flattened version of their opponents and may leave with more distorted views than before. A real human conversation partner, by contrast, is not a caricature. They bring complexity and nuance that is difficult to simulate.

The realism problem. Even with guardrails in place, AI conversation partners lack realism in ways that matter for dialogue education. Constructive dialogue is about conversations between *people*. AI has no genuine stakes, emotions, or lived experience. Early research on Conversate, a tool for simulated interviews, found that participants viewed the AI as failing to capture the real pressure of a job interview.[31] Unlike a coaching exercise, where AI's limitations can be designed around, conversation partner practice asks AI to be something it cannot be: a person for whom the conversation genuinely matters.

Data exposure. All three AI roles generate sensitive records of students' political and moral beliefs, but the risk is most acute with AI conversation partners, where students may express views they wouldn't share with another person. We address data governance in the deployment section below.

Under the current evidence, AI conversation partners have not yet shown clear skill-building benefits, but carry serious risks. As such, caution is advised. Realizing the potential of this role will require design solutions that current tools have not yet demonstrated—particularly around constraining persuasive influence, presenting opposing views with appropriate realism and nuance, and ensuring skill transfer. It's crucial that researchers and educators make progress here, as the default alternative for students isn't no AI; it's engaging with generic chatbots that have limited safeguards.

Deploying AI for Constructive Dialogue Responsibly

Generative AI is already a part of students' lives. A 2026 Coursera survey of over 4,000 students and faculty across five countries found that 95% are already using AI in educational contexts, yet only 26% of institutions have a formal policy governing that use.[5] Students are already using AI tools to explore ideas, rehearse arguments, and navigate disagreement without pedagogical design, institutional oversight, or any of the guardrails described above. For campus leaders, the question is not whether AI will shape the way students engage across differences, but whether *institutions* will shape how that happens.

Below, we provide guidance for campuses seeking to leverage the scalability of AI for dialogue skill building.

1. **Begin with AI coaching, where the evidence is strongest and the risks are most manageable, and expand deliberately from there.** AI mediation is a possible next step for campuses ready to support structured conversation across differences. AI conversation partners, under current evidence, are not yet a recommended path. Well-facilitated human dialogue remains the stronger option for addressing the exposure barrier. Across all three roles, attention to how these tools function for students from different backgrounds is a cross-cutting requirement. Coaching rubrics, mediation protocols, and conversation partner personas all require careful development and review to ensure they function reliably for a wide range of students.
2. **Be aware that the evidence base is still emerging.** Most existing studies measure immediate or short-term improvements in skills within structured exercises. Much less is known about whether those skills transfer to real-world conversations. This question of transfer remains a central research challenge in the field, and campuses should approach early AI implementations with a learning mindset.
3. **Opt for tools designed with pedagogical guardrails.** This is where purpose-built dialogue tools have a decisive advantage over generic AI. A student using an off-the-shelf LLM has no guarantee of pedagogical structure or institutional oversight of their data. Purpose-built AI tools can address these concerns when designed correctly.
4. **Taking no action is itself a risky course.** The reality is that students are already engaging with AI daily. When institutions decline to engage, they do not prevent students from using these tools—they simply relinquish any influence over how they are used. Students will still turn to AI to rehearse arguments, test ideas, and explore political questions, but they will do so with systems that lack the appropriate guardrails. In practice, the decision is no longer whether to adopt AI on campus, but how our institutions will shape its use.

Conclusion

How can generative AI support constructive dialogue on college campuses? The evidence reviewed here points to a clear pattern. AI can most meaningfully support constructive dialogue when its role is constrained and pedagogically structured. As AI tools inhabit a role that is less constrained—from coach to mediator to conversation partner—the risks shift in nature and grow more significant. For coaches, the risks are primarily pedagogical: whether the tool is designed well enough to build the skills it promises. For mediators, they extend to how AI handles facts and manages live disagreement between people. For conversation partners, they reach their most critical form: the capacity to shift beliefs and distort understanding of those we disagree with. AI coaching offers real promise for skill-building at scale, and AI mediation offers a potentially promising path to structured conversations across differences. AI conversation partners, under the current evidence, carry risks that outweigh their benefits until further progress is made.

These conclusions matter beyond college campuses, as does the choice campus leaders make about how to respond to them. Language is indispensable to democratic life, and AI now generates billions of words daily, meaning that a growing share of the language shaping public discourse is produced by algorithms rather than citizens. At the same time, students are already using AI, with or without institutional guidance. Campus leaders who engage thoughtfully with AI in dialogue skill-building are not just serving their students, they're bringing pedagogical and civic intention to a technological revolution that urgently needs it.

The skills of constructive dialogue can be practiced with AI. But the deeper capacities that make dialogue transformative (intellectual humility, the moral courage to voice an unpopular truth, and the vulnerability of admitting uncertainty to someone you disagree with) develop primarily through sustained human experience, not through AI feedback alone. The relational and developmental work that transforms how students engage across differences remains fundamentally human. Yet AI, used wisely as a supplement to proven human-facilitated approaches, can help students build the mechanics, find confidence through early success, and arrive at those human encounters better prepared to make them count. The task for campus leaders is to act on both of those truths at once.

Bibliography

1. Pew Research Center, “The Political Values of Harris and Trump Supporters,” August 26, 2024, https://www.pewresearch.org/politics/2024/08/26/the-political-values-of-harris-and-trump-supporters.

2. Pew Research Center, “As Democrats’ Anger Spikes, Americans’ Feelings about the Federal Government Grow More Polarized,” December 4, 2025, https://www.pewresearch.org/short-reads/2025/12/04/as-democrats-anger-spikes-americans-feelings-about-the-federal-government-grow-more-polarized.

3. Pew Research Center, “Most Americans Say Republican and Democratic Voters Cannot Agree on Basic Facts,” July 30, 2025, https://www.pewresearch.org/short-reads/2025/07/30/most-americans-say-republican-and-democratic-voters-cannot-agree-on-basic-facts.

4. OpenAI, “Ensuring AI Use in Education Leads to Opportunity,” March 5, 2026, https://openai.com/index/ai-education-opportunity.

5. Marni Baker Stein, “AI in Higher Education Report 2026,” *Coursera*, February 25, 2026, https://www.coursera.org/enterprise/resources/ebooks/ai-in-higher-ed.

6. Constructive Dialogue Institute, “What Is Constructive Dialogue?” https://constructivedialogue.org/what-is-constructive-dialogue.

7. William J. Brady, Killian McLoughlin, Tuan N. Doan, and Molly J. Crockett, “How Social Learning Amplifies Moral Outrage Expression in Online Social Networks,” *Science Advances* 7, no. 33 (2021): eabe5641, https://doi.org/10.1126/sciadv.abe5641.

8. Miller McPherson, Lynn Smith-Lovin, and James M. Cook, “Birds of a Feather: Homophily in Social Networks,” *Annual Review of Sociology* 27, no. 1 (2001): 415–444, https://doi.org/10.1146/annurev.soc.27.1.415.

9. Erik Santoro and Hazel Rose Markus, “Listening to Bridge Societal Divides,” Current Opinion in Psychology 54 (2023): 101696, https://doi.org/10.1016/j.copsyc.2023.101696.

10. Martin Saveski, Nabeel Gillani, Ann Yuan, Prashanth Vijayaraghavan, and Deb Roy, “Perspective-Taking to Reduce Affective Polarization on Social Media,” In *Proceedings of the International AAAI Conference on Web and Social Media* 16 (2022): 885–895, https://doi.org/10.1609/icwsm.v16i1.19343.

11. Douglas Stone, Bruce Patton, and Sheila Heen, *Difficult Conversations: How to Discuss What Matters Most* (Viking, 1999).

12. Jonah Koetke, Karina Schumann, Keith Welker, and Peter T. Coleman, “Intellectual Humility Is Reliably Associated with Constructive Responses to Conflict,” *PLOS ONE* 19, no. 9 (2024), https://doi.org/10.1371/journal.pone.0309848.

13. Eran Halperin, “Emotion, Emotion Regulation, and Conflict Resolution,” *Emotion Review* 6, no. 1 (2013): 68–76, https://doi.org/10.1177/1754073913491844.

14. K. Anders Ericsson, “Deliberate Practice and Acquisition of Expert Performance: A General Overview,” *Academic Emergency Medicine* 15, no. 11 (2008): 988–994, https://doi.org/10.1111/j.1553-2712.2008.00227.x.

15. Dongliang Ding and Ahmad Muhyiddin B Yusof, “Investigating the Role of AI-Powered Conversation Bots in Enhancing L2 Speaking Skills and Reducing Speaking Anxiety: A Mixed Methods Study,” *Humanities and Social Sciences Communications* 12, no. 1 (2025): 1–16, https://doi.org/10.1057/s41599-025-05550-z.

16. Ashish Sharma, Inna W. Lin, Adam S. Miner, David C. Atkins, and Tim Althoff, “Human–AI Collaboration Enables More Empathic Conversations in Text-Based Peer-to-Peer Mental Health Support,” *Nature Machine Intelligence* 5, no. 1 (2023): 46–57, https://doi.org/10.1038/s42256-022-00593-2.

17. Greg Kestin, Kelly Miller, Anna Klales, Timothy Milbourne, and Gregorio Ponti, “AI Tutoring Outperforms In-Class Active Learning: An RCT Introducing a Novel Research-Based Design in an Authentic Educational Setting,” Scientific Reports 15 (2025): 17458, https://doi.org/10.1038/s41598-025-97652-6.

18. I-Wen Lin, Ashish Sharma, Christopher M. Rytting, Adam S. Miner, Jina Suh, and Tim Althoff, “IMBUE: Improving Interpersonal Effectiveness through Simulation and Just-in-Time Feedback with Human-Language Model Interaction,” arXiv preprint arXiv:2402.12556 (2024), https://arxiv.org/pdf/2402.12556.

19. Mary Aviles and Mylien T. Duong, “Designing Effective Dialogues: What Students Want,” Constructive Dialogue Institute, November 2025, https://constructivedialogue.org/articles/engaging-students-across-difference-guide.

20. Hamsa Bastani, Osbert Bastani, Alp Sungu, Haosen Ge, Özge Kabakcı, and Rei Mariman, “Generative AI without Guardrails Can Harm Learning: Evidence from High School Mathematics,” *Proceedings of the National Academy of Sciences* 122, no. 26 (2025), https://doi.org/10.1073/pnas.2422633122.

21. Lisa Messeri and Molly J. Crockett, “Artificial Intelligence and Illusions of Understanding in Scientific Research,” *Nature* 627, no. 8002 (2024): 49–58, https://doi.org/10.1038/s41586-024-07146-0.

22. Xuechunzi Bai, Angelina Wang, Ilia Sucholutsky, and Thomas L. Griffiths, "Explicitly Unbiased Large Language Models Still Form Biased Associations," *Proceedings of the National Academy of Sciences* 122, no. 8 (2025): e2416228122, https://doi.org/10.1073/pnas.2416228122.

23. Michael H. Tessler, Max A. Bakker, Daniel Jarrett, Hannah Sheahan, Martin J. Chadwick, Raphael Koster, and Christopher Summerfield, "AI Can Help Humans Find Common Ground in Democratic Deliberation," *Science* 386, no. 6719 (2024): eadq2852, https://doi.org/10.1126/science.adq2852.

24. SwayBeta. "Early Findings with Experimental Participants," Accessed February 20, 2026, https://www.swaybeta.ai/research/experimental-data.

25. Hause Lin, Gabriela Czarnek, Benjamin Lewis, Joshua P. White, Adam J. Berinsky, Thomas Costello, Gordon Pennycook, and David G. Rand, "Persuading Voters Using Human–Artificial Intelligence Dialogues," *Nature* 648 (2025): 394–401, https://doi.org/10.1038/s41586-025-09771-9.

26. Thomas H. Costello, Gordon Pennycook, and David G. Rand, "Durably Reducing Conspiracy Beliefs through Dialogues with AI," *Science* 385, no. 6714 (2024): eadq1814, https://doi.org/10.1126/science.adq1814.

27. Thomas H. Costello, Kellin Pelrine, Matthew Kowal, Antonio A. Arechar, Jean-François Godbout, Adam Gleave, David Rand, et al., "Large Language Models Can Effectively Convince People to Believe Conspiracies," arXiv preprint arXiv:2601.05050 (2026), https://doi.org/10.48550/arXiv.2601.05050.

28. Philipp Schoenegger, Francesco Salvi, Jiacheng Liu, Xiaoli Nan, Ramit Debnath, Barbara Fasolo, Evelina Leivada, et al., "Large Language Models Are More Persuasive than Incentivized Human Persuaders," arXiv preprint arXiv:2505.09662 (2025), https://doi.org/10.48550/arXiv.2505.09662.

29. Ryan Shea, Aymen Kallala, Xin Lucy Liu, Michael W. Morris, and Zhou Yu, "ACE: A LLM-based Negotiation Coaching System," in Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing (2024): 12720–12749, https://doi.org/10.18653/v1/2024.emnlp-main.709.

30. Jeffrey Lowenstein and Leigh Thompson, "Learning to Negotiate: Novice and Experienced Negotiators," In *Negotiation Theory and Research*, ed. Leigh Thompson (Psychology Press, 2006), 77–97.

31. Tiffany Daryanto, Xi Ding, Laura T. Wilhelm, Sabrina Stil, Kristopher M. Knutsen, and Eugenia H. Rho, "Conversate: Supporting Reflective Learning in Interview Practice through Interactive Simulation and Dialogic Feedback," *Proceedings of the ACM on Human-Computer Interaction* 9, no. 1 (2025): 1–32, https://doi.org/10.1145/3701188.