Where Are the Citations?: ChatGPT Discussions in the History of Sexuality

Chris Babits
Utah State University

During the spring 2023 semester, I experimented with ChatGPT in my course on the history of sexuality in the United States. Several students had fallen behind in this asynchronous online class. Because I recognize that this course delivery method can present unique challenges for online learners, I wanted to provide these students the opportunity to earn a strong grade.[1] I also wanted to see how students engaged with generative artificial intelligence (AI). For these reasons, I created four bonus discussion posts where students could use ChatGPT to advance their knowledge of the history of sexuality. These assignments, which I simply titled "ChatGPT and the History of Sexuality," challenged students to critically engage with the information provided by generative AI.

I built my course on the history of sexual on a zero-based grading system, an assessment strategy that makes it possible for the instructor to add assignment options late in the semester.[2] In this approach, every student starts with zero points. Each time the student completes an assignment, though, they earn points toward their final grade. Students in history of sexuality had a range of assignments they could complete for points, from weekly discussion posts and lesson plans, critical analyses of historical monographs, and a research paper. (See Figure 1: Assignment Breakdown for more information on the course's assignment choices.) Students could complete any number of these assignments to build the 94 points required for an A. As I have written elsewhere, the zero-based grading system, with many assignment choices in the course, makes students active agents over their own learning.[3]

After combing through the grade book in the tenth week of the semester, I discovered that although some students fell behind on earning points, most students were approaching the course in the same way. They completed the weekly discussions, which asked for them to select and analyze a piece of evidence from a lecture or the required reading, and they turned in one or two of the critical analysis papers. To earn points, nearly every student submitted at least one lesson plan, an option I include for the teaching majors and minors who enroll in my courses. Ultimately, only three students turned in a research paper, which had to be eight-to-ten pages in length. Even with all these assignment choices, a handful of students found themselves short of points as we approached the semester's end. Instead of asking students to go back and complete assignments for which due dates had passed, I created another kind of discussion board. In these discussions, students engaged with ChatGPT.

For the last four weeks of the semester, students could turn in a discussion post where they were responsible for three tasks. First, the assignment required students to type into ChatGPT a question that the week's material

[1] See the following articles on how students can struggle in asynchronous online learning environments: Doris U. Bollinger and Florence Martin, "Instructor and Student Perceptions of Online Student Engagement Strategies," *Distance Education* 39, no. 4 (2018): 568–583; Amber D. Dumford and Angie L. Miller, "Online Learning in Higher Education: Exploring Advantages and Disadvantages for Engagement," *Journal of Computing in Higher Education* 30, no. 3 (2018): 452–465; Jenna Gillett-Swan, "The Challenges of Online Learning: Supporting and Engaging the Isolated Learner," *Journal of Learning Design* 10, no. 1 (2017): 20–30; and Florence Martin, Chuang Wang, and Ayesha Sadaf, "Student Perception of Helpfulness of Facilitation Strategies That Enhance Instructor Presence, Connectedness, Engagement, and Learning in Online Courses," *The Internet and Higher Education* 37 (2018): 52–65.

[2] Delphi Center for Teaching and Learning, "0-Based Grading Concepts," *Delphi Deep Dive: Design Cookbooks*, accessed October 9, 2024, https://tutorials.delphi-me.com/deep_dive/design-cookbooks/0-vs-based-grading/0-based-grading-concepts.

[3] Chris Babits, "'I Can Do What I Want?' Student Agency in a U.S. History Survey," *The History Teacher* 57, no. 3 (May 2024): 329-355.

Teaching History, 49(1)
DOI: 10.33043/a2az53dz2

(i.e., lectures and readings) helped to answer. In the assignment directions, I asked students to open their post by letting the class know the question they asked ChatGPT. Second, students had to copy-and-paste the answer ChatGPT gave them into their discussion post. For this, I wanted everyone in the class to see the exact words ChatGPT gave in its response. Last, students had to analyze ChatGPT's answer. For this part of the assignment, they needed to share the strengths and shortcomings of the answer ChatGPT provided them. I asked, *Was the answer adequate?* This part of the assignment, which is what I was most interested in, required students to write six-to-eight sentences where they analyzed ChatGPT's answer.

If students completed these parts of the ChatGPT and the History of Sexuality assignment, they could earn up to five points toward their final grade. Students could also earn one or two bonus points if they advanced the discussion. To earn more than five points on the ChatGPT discussion, students could build on their classmates' insights into the answer ChatGPT offered. They could also note other strengths or shortcomings in the answer ChatGPT provided. Or students could share their general thoughts on AI bots after reading some of their classmates' posts.

From the moment I conceptualized this assignment, I hoped that by keeping responses somewhat flexible, students would have multiple entry points into participating in the discussion. I thought that these discussion posts could reinforce some of the learning goals I have in my courses, such as asking students to explain the important of the history of sexuality in their own words while applying historical methods to analyze sources. Most importantly, I want students to ask and develop questions that interest them. Doing so, I believe, challenges students to take control over their own learning. As Johnson notes, students can construct deeper meanings of online discussions via personal contributions, particularly through collaboration and knowledge construction.[4] In addition, I want students to be skeptical consumers of texts and sources. As a historian, I believe that every source has a purpose and perspective to share. But what purpose and perspective does generative AI have? Though there is no definitive answer to this question at the moment, my course's ChatGPT discussion board could provide some insights on the question.

Students could complete the ChatGPT and the History of Sexuality assignment for the first time during Week #12. This week focuses on the sexual revolution, with lectures titled "Women, Sex, and the Pill," "Sex and the Liberated Woman," and "Liberation." (This last lecture covers the early history of the gay rights movement.) The required readings offer a different perspective on the sexual revolution. Readings include primary and secondary sources on Paul R. Ehrlich's *The Population Bomb*, a book that argued that humanity would face destructive famines and hunger if the world's population continued to grow. Week #12 offered students a wide range of material around which they could ask ChatGPT a question.

After Week #12's posts were due, I was pleased with my students' ability to pose strong questions to ChatGPT. More specifically, students asked ChatGPT penetrating questions that demonstrated how closely they watched lectures and engaged with the required reading. Students wrote excellent questions, for instance, about Helen Gurley Brown's *Sex and the Single Girl*, the birth control pill, Anne Koedt on the myth of the vaginal orgasm, Stonewall, and how the sexual revolution changed life for African American women. These were all topics that lectures had covered to some extent. But it was clear students wanted to know even more about these parts of the sexual revolution.

Although I was impressed by the questions students asked, I was less enthusiastic about how students interpreted the answers ChatGPT provided. Put simply, ChatGPT generated what I thought were facile historical understandings of the sexual revolution. In its best form, ChatGPT constructed three or four paragraphs about events like the Stonewall Riots. In the worst of the answers, ChatGPT offered vague numbered lists about something the sexual revolution changed in the United States. No matter the format of ChatGPT's answer, though, students seemed generally impressed with the generative AI's ability to respond to their questions. One student even doubted if *any* human could ever write a more detailed answer than what ChatGPT offered! (In

[4] Cynthia M. Johnson, "Rethinking Online Discourse: Improving Learning through Discussions in the Online Classroom," *Education and Information Technologies* 21, no. 6 (2016): 1483–1507.

my assignment comment to this student, I assured them that though I could not craft an answer as quickly as generative AI, I was certain that my many years of formal study of the history of sexual ensured that my response would be far more detailed and nuanced.) For the first week I offered this assignment, the most critical comments revolved around generative AI's tendency to be indistinct. But that was it. My impression at this point was the students were largely impressed that this new tool could create an answer to their question in such a short span of time.

Something then happened the following week that changed the apprehension I felt with this discussion post assignment. In Week #13, the course moved into histories of sexuality in the 1970s and 1980s. Students watched lectures on the sexualization of American culture and the Religious Right's backlash to changing gender and sexual norms. They also read primary and secondary sources on marketing safe sex and the AIDS crisis. The week's content on the Religious Right lit a fire with my students, many of whom identify as lesbian, gay, bisexual, queer, trans, nonbinary, questioning, or a strong ally. As such, I saw much more critical engagement with the answers ChatGPT generated to their questions.

Students noted that ChatGPT made several inaccuracies in its answers about the content we covered in Week #13. In their posts, students also explained how they received lackluster answers to their initial questions. Because of this, I started to task my students with being prompt engineers, challenging them to refine the questions they asked the generative AI. A good number of students were still impressed by ChatGPT, for sure. But Week #13 offered a first for this assignment—a student asking where the generative AI culled its information from! This student wanted citations. They were concerned that without sources, there would be no way to verify the information they were reading. This aspect of ChatGPT troubled them.

Over the next two weeks, students continued to refine their critical analyses of ChatGPT's answers. A greater number of students started to wonder where the citations were. Others realized that they could ask ChatGPT to insert citations, though they were not sure if the citations were strong ones. Even worse, they observed, was how ChatGPT would not include accurate direct citations from primary sources. My students also started to muse about the future of education in a world with generative AI. They noted that ChatGPT offered well-written answers to a range of questions. But they also acknowledged how leaden and boring the AI's prose was. I could not have agreed more. After reading these more critical responses to ChatGPT, I literally belted out, "The kids are alright!"

My experience using ChatGPT in my course on the history of sexuality showed me that there are uses for generative AI. Rather than lambasting technological change, as many educators have since OpenAI released ChatGPT into the world, I see the future as one of pedagogical possibilities. Of course, we should be concerned that students may shortchange their own learning by turning to generative AI to complete their work for them. This might be especially true in introductory courses where we want students to build a strong foundation for college success and career readiness. This represents a clear tension that we may all feel in our introductory surveys.

But we cannot bury our heads in the sand because we do not like new technologies. Instead, as my course's ChatGPT discussion showed me, we must introduce our students to tools like generative AI. Doing so will prepare them for the world they encounter outside our classrooms. Not teaching our students how to use these tools—and that we know about generative AI's power and promise—will only lead to tension with our students, who will rightly say that we are living in the past.

References

Babits, Chris. "'I Can Do What I Want?' Student Agency in a U.S. History Survey." *The History Teacher* 57, no. 3 (May 2024): 329-355.

Bollinger, D. U., and Florence Martin. "Instructor and Student Perceptions of Online Student Engagement Strategies." *Distance Education* 39, no. 4 (2018): 568–583. https://doi.org/10.1080/01587919.2018.1520041.

Delphi Center for Teaching and Learning. "0-Based Grading Concepts." *Delphi Deep Dive: Design Cookbooks*. Accessed October 9, 2024. https://tutorials.delphi-me.com/deep_dive/design-cookbooks/0-vs-based-grading/0-based-grading-concepts.

Dumford, Amber D., and Angie L. Miller. "Online Learning in Higher Education: Exploring Advantages and Disadvantages for Engagement." *Journal of Computing in Higher Education* 30, no. 3 (2018): 452–465. https://doi.org/10.1007/s12528-018-9179-z.

Gillett-Swan, Jenna. "The Challenges of Online Learning: Supporting and Engaging the Isolated Learner." *Journal of Learning Design* 10, no. 1 (2017): 20–30. https://doi.org/10.5204/jld.v9i3.293.

Johnson, Cynthia M. "Rethinking Online Discourse: Improving Learning through Discussions in the Online Classroom." *Education and Information Technologies* 21, no. 6 (2016): 1483–1507. https://doi.org/10.1007/s10639-015-9395-3.

Martin, Florence, Chuang Wang, and Ayesha Sadaf. "Student Perception of Helpfulness of Facilitation Strategies That Enhance Instructor Presence, Connectedness, Engagement, and Learning in Online Courses." *The Internet and Higher Education* 37 (2018): 52–65. https://doi.org/10.5204/jld.v9i3.293.

Figure 1: Assignment Breakdown

Assignment Name	Short Assignment Description	Frequency of Assignment	Points Possible from Assignment: Per Assignment / Total for the Course
Commonplace Journal Entries	A discussion post where students select a piece of evidence from a lecture or reading, analyze that evidence, and respond to two of their classmates.	One per week, for the semester's fourteen weeks.	5 points / 70 points
Lesson Plans	A mock version of creating a lesson plan, where students turn the content they learned and transform it into an age-appropriate lesson plan for high schoolers.	One per week, for the semester's fourteen weeks.	5 points / 70 points
Critical Analyses	A short paper (4-to-5-pages) where students critical evaluate the strengths and shortcomings of one of the course's required historical monographs.	Four total, one for each of the course's required monographs.	12 points / 48 points
Research Paper	An eight-to-ten-page research paper on one of the following topics: rape and sexual power in colonial America; the masculinity of the enslaved in the antebellum era; miscegenation law and the making of race in America; abortion before *Roe v. Wade*, and homosexuality and the *Diagnostic and Statistical Manual.*	One.	30 points
ChatGPT and the History of Sexuality	An online discussion post where students ask ChatGPT a question based on the week's material, copy-and-paste ChatGPT's response to their question, and offer six-to-eight sentences of analysis about the generative AI's answer.	Four assignments, starting with Week #12, continuing with Week #13 and Week #14, and an opportunity to submit again during finals.	5 points / 20 points
Extra Credit for ChatGPT and the History of Sexuality	Respond to one or two comments from classmates.	Four opportunities. See above for more on assignment frequency.	Up to 2 points / Up to 10 points