

# “It’s Giving AI”: Reading Ambiguously-Authored Texts and the Role of Felt Sense

Elizabeth Velasquez<sup>1</sup> & Christa Teston<sup>1</sup>

<sup>1</sup> Ohio State University | US

**Abstract:** To understand how human readers navigate a literate landscape that newly includes AI-generated prose, we asked participants (n=76) to read and make decisions about who and/or what is responsible for writing anonymized, “ambiguously-authored” texts. Findings suggest that readers’ assumptions about who and/or what wrote a text are rooted in “felt sense.” Prompting participants to make their “felt sense” explicit allowed us to catalog the evidential warrants participants relied on when making authorship decisions. Enabled by a modified grounded theory approach to analysis, we constructed two main themes. First, readers are “triggered” by certain textual cues that, when combined with prior experiences and knowledge, evidentially warrant assumptions about who and/or what wrote a text. Second, after recognizing the consequences of making one’s felt sense explicit, some readers experience what we call an “axiological crisis.” Axiological crises emerge when participants meta-cognitively hear or see themselves attributing certain characteristics and values to an AI text-generator or human author. We conclude by reimagining the axiological crisis as an opportunity for improving metacognitive awareness about how felt sense affects our reading practices.

**Keywords:** tacit knowledge, meaning-making, RX Lab, evidence, grounded theory



Velasques, E., & Teston, C. (2026). “It’s giving AI”: Reading ambiguously-authored texts and role of felt sense. *Journal of Writing Research*, 17(3), 581-619. DOI: <https://doi.org/10.17239/jowr-2026.17.03.08>

Contact: Elisabeth Velasquez, Ohio State University | US - [velasquez.85@buckeyemail.osu.edu](mailto:velasquez.85@buckeyemail.osu.edu)

Copyright: This article is published under Creative Commons Attribution-Noncommercial-No Derivative Works 3.0 Unported license.

Now that generative artificial intelligence (AI) is part of our communicative ecosystem, as writing instructors, we find ourselves constantly curious (if not cynical) about with whom and/or what we're communicating. This problem is not unique to educational contexts, though. Motivated by a desire to reduce how frequently readers might be deceived by "synthetic" prose (Valiaiev, p. 3), experts in AI have studied how well human and nonhuman readers are able to discern differences between human-authored vs. AI-generated prose. When it comes to online shopping, for example, Adelani, et al. (2020) concluded that human (n=80) and nonhuman readers are "far from perfect" at detecting fake reviews (p. 11). To help improve human readers' "skill at detecting machine-generated" prose, Dugan, Ippolito, Kirubarajan, & Callison-Burch (2020) designed a "Real or Fake Text" website that simultaneously collects data from human readers about what they perceive as being more or less human while gamifying the practice of detecting AI prose within a corpus of short stories, recipes, news articles, and presidential speeches (p. 1). Sensitive to how "real or fake text" detection might hinge on disciplinary expertise, Liu, et al. (2024) compare how well medical professionals (n=4) versus "AI detectors" (e.g., Originality.ai, TurnItIn, GPTZero, ZeroGPT, etc.) can identify AI-generated prose within medical writing. Ultimately, they argue in favor of incorporating "AI content detectors...as an additional screening tool in the peer-review process of academic journals" (n.p.). These are but a small sampling of impressive quantitative studies that are both internally and externally valid approaches to understanding the problem of reading ambiguously-authored texts.

By characterizing a text as "ambiguously-authored," we are drawing attention to the fact that participants in this study were explicitly told that the anonymized abstracts they read may have been written by (a) a human, (b) AI, or (c) human + AI. Readers must then navigate uncertainty about a text's authorial origins by reasoning out loud about why they assume a particular text had been written by a human and/or AI. To capture participants' reasoning, we relied on qualitative (albeit messy) methods, such as surveys, talk-aloud protocols, focus groups, and semi-structured interviews.

Our project contributes to ongoing conversations in AI by offering a qualitative approach to capturing human readers' evidential processes associated with reading ambiguously-authored prose. We're curious about the perhaps unquantifiable kinds and forms of evidence that readers draw on when assessing written communication's authorship. Our humanistic approach to the problem of contemporary reading does not begin with a hypothesis about how (in)capable humans are at detecting AI-generated prose. Nor do we begin with *a priori* definitions for what counts as evidence for detecting AI-generated prose.<sup>1</sup> And unlike some colleagues in writing studies whose work leans more toward AI refusal (Sano-Franchini, McIntyre, Fernandes, 2025), our study design is AI-agnostic. In what follows, we present

---

<sup>1</sup> Inspired by Walters 2019 and Lee 2022, Liu, et al. (2024) asked human participants to deductively "select their top three reasons" for their authorship decisions based on a list of prepared options; these included: "coherence creativity, evidence-based, grammatical errors, and vocabulary diversity" (p. 5). Unlike Liu, et al., the evidence we collect in our study is inductively derived.

results of a qualitative, empirical investigation of human readers’ (n=76) process of deciding whether an anonymized text was written by humans and/or AI and what led them to that conclusion (IRB #2024E1056).<sup>2</sup>

This project also contributes to ongoing conversations about AI in writing studies—studies that have demonstrated how important it is to employ messy, humanistic approaches to studying the problem of reading ambiguously-authored texts. Writing studies research on AI tends to offer practical implications for, say, teaching professional (Bedington, Halcomb, and McKee, 2024), first-year (Cummings, Monroe, Watson, 2024), or multimodal writing (Burris, Smith, Shimizu, Hundley, Pendergrass, 2025). Although findings from such studies may not be generalizable, the richly descriptive constructs and historical insights that emerge from writing studies research into human/machine writing may be operationalized by colleagues beyond our discipline (see Laquintano, Schnitzler, and Vee’s [2023] “A Brief History of Automated and Computational Writing”). For example, Knowles (2024) recasts the binary of “human vs. nonhuman” writers as “rhetorical load sharing.” That is, authorship might be theorized as spanning a spectrum “from human-authored text to synthetic text” (n.p.). Pandey, Bhusal, and Niraula (2024) likewise offer nuance to contemporary discussions about large language models in the classroom by framing them in terms of “encroachment” or “enhancement.” Our study continues the rich tradition of thick description in writing studies research by offering a methodologically transparent representation of readers’ **felt sense** when reading ambiguously-authored texts.

As an intellectual anchor for what we mean by “reading,” throughout this essay, we rely on Haas and Flower (1988) who describe reading as “a process of responding to cues in the text and in the reader’s context to build a complex, multi-faceted representation of meaning” (p. 169). Our analyses zoom in on participants’ verbal and textual utterances about the “cues in the text” that participants see as evidentially significant when determining a text’s authorship. Importantly, participants tether said “cues in the text” with a range of prior experiences and knowledge when deciding whether prose is “real” or “synthetic” (Valiaiev, p. 3).

What does a participant mean when they say, as our title suggests, that a text “is giving AI”? In contemporary popular discourse, “it’s giving...is usually preceded by a noun/phrase to describe how someone looks or the vibe that they are bringing. For example, someone might describe an outfit they’ve seen someone wear as, ‘it’s giving girlboss’” (BBC, n.p.). Herring (2024) points out that “it’s giving” is rooted in “queer vernacular” and has subsequently entered our contemporary lexicon thanks to “shows such as *RuPaul’s Drag Race* and *Queer Eye*” (p. 260). When one of our participants declared that the text we’d asked them to read was “giving AI,” we were inspired to look closer at how AI-generated prose has a kind of “vibe.” We wondered if “it’s giving” might be TikTok-speak for felt sense, in other words. Ultimately,

---

<sup>2</sup> In adherence to JOWR’s commitment to data availability and open science, the anonymized dataset, code for data analysis, and any other research materials will be made available through Ohio State University’s Reader Experience Lab’s website ([go.osu.edu/RXLab](http://go.osu.edu/RXLab)).

this project sheds light on how participants draw on and describe their “felt sense” (Perl 2004; Polanyi 1967) while reading ambiguously-authored texts.

### 1. Some Notes on “Felt Sense”

“Something about the language pattern. It’s hard to say exactly what tipped me off.”

Asynchronous cohort participant

Writing studies scholars have productively mobilized “felt sense” and/or “tacit knowledge” in their research by drawing attention to, for example: writers’ style (Olinger, 2014); writers’ transferable skills (Adler-Kassner, Clark, Robertson, Taczak, & Yancey, 2016 [see threshold concept #5]); writers’ “para-expertise” (Rice, p. 119); writing educators’ practices (Crozier & Workman, 2022; Fleckenstein, 2003; Kim & Cho, 2017; Rule 2017; Wolfe 2005); writers’ processes (Perl 1979, 1980, 2004); and how professional writers collaborate and communicate about risk and/or uncertainty across differences (Clayson 2018; Evia & Patriarca, 2012; Haas & Witte 2001; Sauer 2003). Importantly, some writing studies scholars have helped to unpack the nuances in, for example, Collins’s tacit knowledge taxonomy (Amidon 2022). But as explanatory constructs, felt sense and tacit knowledge long precede writing studies.

In this essay, we hope to inductively derive a local account of felt sense, inspired by the uncertainty that the existence of AI-generated prose presents. Sondra Perl’s understanding of Eugene Gendlin’s origination of “felt sense” is that it indexes “an unclear, barely noticeable bodily sensation” that can sometimes be “slightly disturbing because it calls attention to what is just on the edge of our thinking but not yet articulated in words” (p. xiii). Perl goes on to describe felt sense as “inchoate pushes and pulls,” or “barely formed preverbal...leanings” (p. xiii). Of course, writing studies scholars’ use of “felt sense” is also inspired by Michael Polanyi’s (1967) “tacit knowledge.” In his book, *The Tacit Dimension* (p. xiii), Polanyi argues that we “know more than we can tell” (p. 4). Prior to publishing *The Tacit Dimension*, Polanyi (1958) emphasizes how incredibly personal and “ineffable” such knowledge is: “what I call ‘ineffable’ may simply mean something that I know and can describe even less precisely than usual, or even only very vaguely” (Polanyi, p. 88). Similarly, in his foreword to Sondra Perl’s *Felt Sense*, Peter Elbow articulates “felt sense” as “that felt nonverbal sense triggered by a word” (p. v). Later, he goes on to say that “when we hear the ‘offness’ in a word we’ve used, we hear it even if it’s just *slightly off*” (p. ix; emphasis in the original). Regardless of who or what actually authored a text, now that AI is part of our literate landscape, readers may be highly attuned to “offness.” For us, sensing an “offness” indicates the presence of felt sense.<sup>3</sup>

Extrapolating from writing studies scholarship on felt sense, and inspired by scholars’ admonition that writing studies must better account for writing and reading practices’ entanglement (cf. Carillo; Keller), we wonder about the role of felt sense for readers.

---

<sup>3</sup> For our purposes, we use felt sense interchangeably with tacit knowledge, but we understand that there are good reasons to be more discerning about the boundaries between these two constructs.

Specifically, how do readers yoke “cues in the text” with their own prior knowledge and experience in order to make meaning (Haas & Flower, p. 169)? Ultimately, to understand contemporary literate practices now that AI is in the mix, we must—through fine-grained analyses of real-time reading practices—account for how readers evidentially warrant the assumptions they make about who or what wrote a text.<sup>4</sup> Mining readers’ hunches about who or what wrote the prose they’re reading is a first step toward understanding how, if at all, contemporary reading practices have changed in reaction to AI’s presence. Inductive analyses of survey responses, interview and focus group transcripts, and audio-visual recordings shed light on the textual cues that readers explicitly draw on to arrive at a decision about whether a text has been generated by AI and/or written by a human.

This is not an experimental study that aims to reveal how (un)successful human readers are at detecting when a text is written by an AI text generator. Rather, this is a qualitative, exploratory study that inductively derives one account of the role that “felt sense” plays in readers’ processes now that AI is a part of our literate landscape. For the purposes of this study, we understand “felt sense” to encompass readers’ intuitive, often pre-linguistic “hunches” about a text and its author. Among other factors, these hunches may be inspired by what participants perceive to be evidentially significant “cues in the text” (Haas & Flower, p. 169) and/or their own pre-existing ideas, assumptions, and practices.

## 2. Methodology

We are under no illusions that it is possible to fully apprehend how readers make meaning from and/or ascribe (non)human agency to a text—especially when such processes are, indeed, implicit or pre-linguistic. We concur with Haas and Flower (1988) who rightly argue that “we cannot enter the reader’s heads and watch as the construction of meaning proceeds” (p. 170). Indeed, all we have access to are “indirect measure[s] of the nature, content, and structure of that representation” (p. 170). Although they’re focused on writing practices, Odell, Goswami, and Herrington (1983) similarly confirm the methodological conundrum researchers face when trying to “make explicit the knowledge or strategies that previously may have been only implicit” (p. 223).

One indirect measure for mining readers’ processes includes observing participants, in real time, as they read texts. In addition to real-time observations, we also include semi-structured interview methods and focus groups—not necessarily “to obtain information about mental processes,” but rather “to identify the kinds of world knowledge and expectations” that participants bring to the reading task (Odell, Goswami, and Herrington, p. 228). To account for readers’ “world knowledge and expectations,” we conducted this study in two ways: through

---

<sup>4</sup> Here, we’re invoking Toulmin’s notion of a warrant, which is the argumentative part of speech that logically connects a claim (e.g., about authorship) to the evidence that supports it. Warrants are typically implicit. And importantly, for an argument to be convincing, all interlocutors must be working from the same, agreed upon understanding of the accuracy or believability of a warrant. Our project asks participants to make their evidential warrants *explicit*.

synchronous observation and asynchronous self-reports of how participants read in real time and (explicitly) articulated values-based assumptions about who and/or what wrote study stimuli. The synchronous and asynchronous cohorts of the study had slightly different research designs, which likely affected results. So, after briefly describing the abstracts employed as our study's stimuli, we provide greater detail about the data collection methods we used for each of the cohorts.<sup>5</sup>

## **2.1 About the Study Stimuli: Abstracts from Undergraduate Writers' Published Research**

Because we are situated within an English Department in the United States and teach in one or more of the department's Writing Programs, we are especially interested in how participants warrant their assumptions about what they expected to see within prose that had been generated by AI versus prose that had been written by an *undergraduate writer*. Therefore, study stimuli emerged from anonymized abstracts originally authored by undergraduate writers whose essays were published (prior to the widespread usage of generative AI) in *American Journal of Undergraduate Research*.

After randomly selecting essays from *American Journal of Undergraduate Research*, we redacted all information except for the title and its abstract. Two unmodified abstracts served as the first and second stimuli. The third, fourth, and fifth stimuli were generated by prompting ChatGPT4 and/or Google Gemini to rewrite the abstract. A sixth and final abstract was co-written by both a human and AI (though this stimulus was only deployed during the synchronous cohort). Refer to appendix A for study stimuli, including the prompts we employed when finalizing each stimulus.

## **2.2 Synchronous Cohort Methodological Details**

Prior to data collection sessions, participants were recruited via convenience sampling and asked to complete an online screening survey, hosted in Qualtrics (appendix B). The screening survey gathered information regarding a participant's positionality, prior teaching experience, and familiarity with either using and/or recognizing AI writing tools. Information regarding participants' age and gender was also collected on a voluntary basis. Screening survey results informed how we constructed focus groups of two or three participants. We formed focus groups based on participants' self-identified rank (e.g., graduate students were placed with other graduate students, faculty with other faculty, etc.).

Prior to engaging with the abstracts, focus group participants were asked a series of semi-structured interview questions about their experience with generative AI in the classroom, including whether they implemented policies about student use of AI (pre-test interview questions can be found in appendix C). Participants were then given hardcopies of three social

---

<sup>5</sup> Before executing the full study, we conducted a small-scale pilot study in October 2024 (n=4). Further details regarding the pilot study protocol and findings can be located on the website of the Ohio State University's Reader Experience Lab (link: [go.osu.edu/RXLab](http://go.osu.edu/RXLab)).

scientific abstracts authored by (1) an undergraduate student; (2) an undergraduate student+AI; and (3) ChatGPT4 or Google Gemini. We informed participants of all three potential authors (human, human + AI, and AI), and then asked participants to articulate who/what they thought had authored each abstract, noting that there was one of each. Focus group participants completed the task collaboratively while talking aloud with one another.

After participants completed their task, they were separated so that individual, semi-structured interviews could be conducted in a way that participants felt as though they could speak freely (consult appendix D for the post-test interview questions). Interview questions invited participants to (re)articulate the assumptions they made about authorship and, when necessary, we prompted participants to further explain their reasoning. Finally, using a scale of 1-10 (1=least confident; 10=most confident), we asked participants to indicate how confident they felt about their authorship decisions. After noticing that participants frequently left annotations and jottings on the hard copies of the abstracts, we also included them in our data corpus (figure 1).

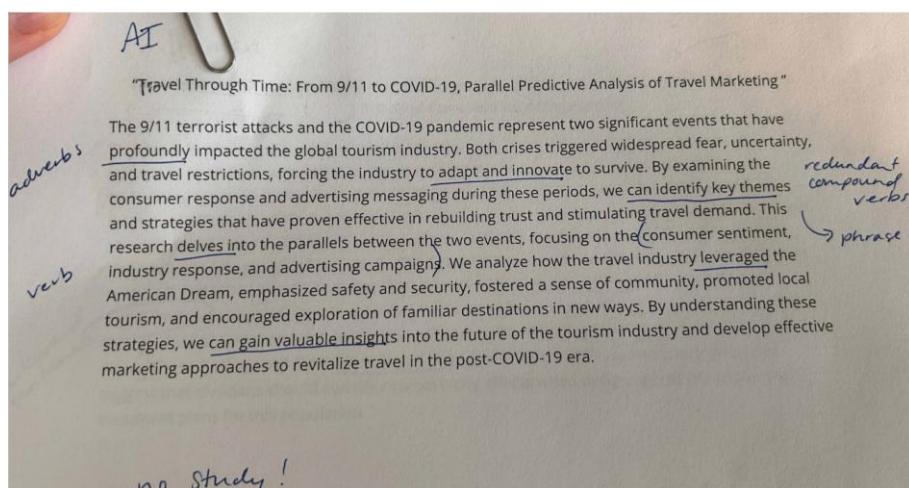


Figure 1. Image of study stimulus with a participant's annotations and jottings

The synchronous cohort consisted of six focus groups, or 15 participants. Across these six focus groups, participants included:

- eight faculty members with teaching experience,
- two members of university staff with teaching experience,
- one graduate student who was also a member of university staff (without teaching experience), and
- four graduate students with teaching experience.

Three focus groups (eight participants) received an abstract written by ChatGPT4, and three (seven participants) received an abstract written by Google Gemini. As we describe elsewhere (Velasquez & Teston, *in press*), these activities took place in the Reader Experience Lab (RX Lab), which is equipped with a one-way mirror observation window and technology that captures real-time audio-visual data (via both a 360° OWL camera and a static camcorder).

In total, we analyzed approximately 400 minutes of audio-visual recordings, 43,000 transcribed words, 15 survey responses, and 15 hard copies of study stimuli from synchronous cohort participants. Although we do not attempt to correlate findings with focus groups' unique disciplinary identities, table 1 details the disciplines represented by the synchronous cohort.

*Table 1.* Disciplinary representation in the synchronous cohort

Focus Group #1	Graduate Teaching Associates (n=2; one in Literature and one in Writing, Rhetoric, and Literacy)
Focus Group #2	Graduate Associates (n=3; two from Literature who have teaching experience and one from the Medical Humanities program who works in the Medical Center and does not teach writing)
Focus Group #3	Staff and faculty from Ohio State University's Center for the Study and Teaching of Writing (n=3; two staff members who have expertise in writing center and/or writing across the curriculum studies; one faculty member with expertise in writing and technical communication)
Focus Group #4	Literature Faculty (n=2)
Focus Group #5	English Department Writing Program Directors (n=3; one staff member who co-directs a writing program; two faculty members who direct different writing programs)
Focus Group #6	Creative Writing Faculty (n=2)

### 2.3 Asynchronous Cohort Methodological Details

In an effort not to limit participation only to people who could meet with us synchronously in the RX Lab, we opted to design an online, asynchronous version of the study (via a Qualtrics survey) that participants could complete on their own time (appendix E includes the asynchronous cohort survey; appendix F includes the abstracts used as study stimuli). For this cohort, we simplified the study to include only two stimuli: an abstract authored by a human

and an abstract generated by AI (ChatGPT4). Unlike the synchronous cohort, the asynchronous cohort’s abstracts were based on the same manuscript authored by an undergraduate student in *the American Journal of Undergraduate Research*.

Using Qualtrics’s “highlight question,” participants were prompted to highlight portions of either abstract as being “indicative of AI generation,” “indicative of human writing,” or “notable for some other reason (explain why).” Participants were prompted to explain their authorship decisions in short-answer format. Finally, we asked participants to rate their confidence in their authorship decisions, using a scale of 1-5 (1=least; 5=most).

The asynchronous data collection cohort included (n= 61<sup>6</sup>):

- 23 tenure-line faculty members;
- 20 undergraduate students;
- ten lecturers, clinical staff, or associated faculty;
- five graduate students (with teaching experience);
- two graduate students (without teaching experience); and
- one participant who declined to identify.

To recruit participants, we relied on convenience and snowball sampling via a variety of professional listservs to which the research team belonged. In total, we analyzed approximately 6,000 words from 61 survey responses from the asynchronous cohort.

### 3. Results and Analysis

“Not to nitpick your study (!), but it’s hard to actually indicate semantically why this feels like AI.”

Asynchronous cohort participant

Although the participant quoted above questions whether the difficulty they’re having with discursively describing why a text “feels like AI” is due to our study’s design, their critique aptly summarizes the project’s most significant finding: It is difficult to describe one’s felt sense about who or what wrote a text. Certainly, “I feel” is often employed as a colloquial assertion rather than an indication that someone is physically feeling something. Yet, given its frequency across the data set, we are confident in asserting that most participants rely on felt sense to make assumptions about the abstracts’ authorship.

A future project will report findings from a gestural analysis of audio-visual recordings; so, in what follows, we stop short of characterizing participants’ reasoning as “embodied,” at least, in the ways that Haas & Witte (2001), Sauer (2003), or even Hayles (1999) might define “embodiment.” Instead, we focus predominantly on what can be empirically apprehended as indicative of someone’s “felt sense.” We agree with the above participant’s critique—i.e., that “we know more than we can tell” (Polanyi, 1958). Still, we see value in identifying felt sense’s linguistic/discursive forms, at least insofar as participants could articulate them, when they

---

<sup>6</sup> Data collection for the asynchronous cohort ran from January 15, 2025 to January 31, 2025.

read each abstract. We wonder how the act of “build[ing] a complex, multi-faceted representation of meaning” (Haas & Flower, p. 169) unfolds when readers are invited to explicitly articulate and then perhaps reflect on the evidential warrants that inform their assumptions about ambiguously-authored texts.

Again, we are less interested in whether and/or how frequently participants accurately identified who/what wrote an abstract; rather, our analyses focus predominantly on the kinds of evidence—i.e., textual cues and hunches—that participants invoked when articulating whether and/or when an abstract is “giving AI.” That said, we understand that our readers are likely curious about how frequently participants were able to distinguish AI-generated prose from human-authored prose. So, in what follows, we summarize general findings from each cohort. Then, we unpack the kinds and forms of evidences readers rely on when making sense of ambiguously-authored texts.

### **3.1 General Findings: Synchronous Cohort**

Within the synchronous cohort (six focus groups,  $n=15$ ), only one focus group (#6) correctly identified who/what wrote all three abstracts. However, no one from the synchronous cohort failed to detect a difference between AI-generated and human-authored abstracts. In other words, no focus group participants mistook human writing for AI writing, or vice-versa. Instead, participants frequently swapped the AI-generated and AI+human co-written abstracts or the human-authored and AI+human co-written abstracts.

To detect potential patterns, we closely examined the four participants who at least initially accurately assumed who/what wrote each abstract. Three of those four participants ultimately ended up compromising with the other members of their focus group, which resulted in inaccurate final decisions about who/what wrote each abstract. When compared with the rest of the focus group participants, these four participants do, indeed, possess greater expertise either in writing pedagogy experience or with using AI. However, the study’s small sample size limits us from making correlations between greater expertise with identifying ambiguously-authored prose and greater experience either with teaching or using AI. In future research, we plan to examine the nature of the three compromising participants’ decision to adopt the will of the group rather than follow their initial hunches. We wonder what evidences the other focus group participants brought up that ultimately convinced these three compromising participants to change their mind and/or abandon their initial felt sense.

No obvious patterns could be correlated to in/accurate assumptions within or across the specific disciplines outlined in table 1. Interview data reveal that none of our participants use an “AI-detection tool,” but a majority of the 14 participants who teach reported that they use some type of AI policy in their classes.

### **3.2 General Findings: Asynchronous Cohort**

Asynchronous cohort participants were presented with only two options (i.e., study stimuli did not include an AI+human co-written abstract), so the likelihood that they would correctly identify who or what wrote an abstract due simply to chance was higher. Since both abstracts

were from the *same* undergraduate's published article, asynchronous cohort participants were able to conduct more direct (i.e., 1:1) comparisons of certain sentence types. One participant commented on how this study design choice may have affected their authorship decision:

This was a tricky one for me, because as I reread these two abstracts over and over again, I first thought that abstract 1 was generated by AI. But then, I began to have my doubts. Thinking it through, I still think abstract 1 was written by AI because of the way I was able to compare it with abstract 2. **If it was on its own, I don't know if I would have been able to distinguish.** (emphasis added)

These study design differences may contribute to the fact that more asynchronous cohort participants (i.e., 65.5 percent, or 40/61 participants) than asynchronous cohort participants correctly identified who/what had written both abstracts.

### 3.3 Evidences that Human Readers Relied On: Toward a Substantive Theory<sup>7</sup>

Each of us independently and then collaboratively coded transcripts, video data, and survey results for the kinds and forms of evidences that participants invoked when articulating decisions about who/what wrote an abstract. Along the way, we also composed and shared analytic memos, coding spreadsheets, and working theories, each of which enabled what grounded theorists might call "constant comparison" (Glaser & Strauss, 1967). The first round of open coding rendered a rather long and unwieldy list of potential codes (Box 1).

---

<sup>7</sup> According to Farkas and Haas (2012), a substantive theory is "a small-t theory concerned with a substantive area," which is simply "one particular situational context" (Strauss & Corbin, 1990, p. 174). Our goal here is to lay the groundwork for replication studies that will help to refine this substantive theory over time.

AI-like_vocabulary	follows_genre_conventions	meta-crisis
appropriate_conclusion	formal	metadiscourse
ascriptions/toAI	general/lacks_specificity	natural
ascriptions/toHuman	general/lacks_specificity	non-repetitive
BottomUp_TopDown	grammatical_finesse	organized
capable_of_caring	grasp_of_vocabulary	parallelism
capable_of_personality	has_ability_to_revise	polished
capable_of_persuasion	hedging	repetition
capable_of_revision	human-like_vocabulary	research_about_AI
capable_of_specificity	idiosyncratic	rhetorical_question
capable_of_writing_process	illogical_conclusions	robotic_pattern
choppy/lacks_connections	impersonal	rumor
clear_argumentation	inability_to_make_errors	sentence_"density"
clear_writing	interest/	sentence_variation
clunky	knowledge_of_this_particular_subject	smooth
clunky/lacks_flow	invention	specific
committee_work	lacks_ability_to_make_	stock_phrases
concise	rhetorical_question	too_fancy
confidence	lacks_confidence	too_general/lacks_specificity
confident	lacks_connections/flow	too_good
confusing	lacks_critical_insight/	uninterested/uninformed
don't_understand_	depth_of_thought	vague/lacks_specificity
genre_conventions	lacks_depth_of_thought	valence
doubt	lacks_flow/connection	varied
excessive_jargon	lacks_grammatical_finesse	well-structured
experience_with_using_AI	lacks_lists	word_repetition
experience_with_grading	lacks_personality	wordy
flow/connections	lacks_revision	
	lacks_vocabulary	
	lists	

Box 1. Open (or first cycle) codes

Although we did not adhere strictly to a traditional grounded theory approach to analyzing our data, the open codes in Box 1 are the result of what Farkas and Haas (2012) term “pushing out / undoing / fracturing” (p. 87; see also Glaser & Strauss, 1967). That is, our first analytic step involved teasing our data apart to get a granular sense of participants’ evidences for determining authorship. Along the way, we began to “provisionally” (p. 87) articulate the properties and dimensions for the evidences we’d coded.

The next significant analytic step involved what Farkas and Haas describe as “pulling in / redoing / building theory” (p. 89), which in traditional grounded theory parlance is known as selective coding.<sup>8</sup> The selective codes that resulted from this process are located in Box 2.<sup>9</sup>

AI experience, committee work, rumor, scholarship  
 characteristics (ascribed to AI; ascribed to human; ascribed to human+AI)  
 decision-paralysis, doubt  
 experience with grading/teaching undergraduate writing  
 flow/connection  
 genre conventions  
 grammar  
 hedging  
 lists  
 meta-crisis, authorship crisis  
 personality  
 punctuation  
 quality of writing  
 sentence structure  
 soundness of argument and/or study  
 style/quality of writing  
 trigger moment  
 valence  
 vocabulary (or word use)

*Box 2. Selective (or second-cycle) codes*

From the selective codes in Box 2, we “pulled in” a bit more to establish what Saldaña (2009) terms “categories” and “themes/concepts.” Table 2 provides more detail about how the selective codes from Box 2 were distilled into four categories. We also share definitions and representative examples for each of the categories.

---

<sup>8</sup> “During selective coding, aspects of the data set that emerged as salient through dimensionalizing are pursued in more detail. The goal...is to identify and to provisionally confirm a core category by systematically relating the core category to data and to other categories” (Farkas & Haas, p. 89).

<sup>9</sup> Due to time constraints, we did not engage in the formal process of inter-rater reliability when coding our data. We independently coded our data, memoed about our preliminary findings, and then met synchronously and asynchronously to discuss next analytic steps.

Table 2. Categories, definitions, selective codes, and representative examples<sup>10</sup>

Categories and Definitions	Selective Codes	Representative Examples
Trigger Definition: A “cue in the text” that helps a participant evidentially warrant a claim about who or what wrote it.	flow / personality / genre conventions / soundness of argument / vocabulary (or word use) / grammar, punctuation, and sentence structure / style	<ul style="list-style-type: none"> <li>“starting with the [rhetorical] question...didn't really feel like the genre.”</li> <li>“too vague”</li> <li>“doesn't connect the dots”</li> <li>“quirky phrasing”</li> <li>“Too slick”</li> <li>“Tone is off”</li> <li>“sentences felt clipped”</li> <li>“Lists of three”</li> <li>“there seems to be this really like, personal element”</li> </ul>
Prior experience or knowledge Definition: A real-world experience, theory, and/or prior knowledge that helps (or would help) a participant evidentially warrant a claim about who or what wrote it.	experience with AI / experience with grading undergraduate writing / AI rumor / AI scholarship / AI committee work	<ul style="list-style-type: none"> <li>“There's this article I just finished reading this afternoon...that talks about...post-human ethics...”</li> <li>“I'm just finishing grading with student drafts and it reads rather similarly to some of them.”</li> <li>“Before I went to grad school, I had a couple of years where I worked on academic misconduct. And at that point it was, it wasn't turnitin, but it was something called safe assign.”</li> <li>“And I have almost no experience reading AI, but I just feel...since these sentences kind of swell with certain information...they have more of a narrative.”</li> <li>“I don't know about writing abstracts. Or if, like, we use personal pronouns in writing abstracts or not.”</li> </ul>
Characteristics and Attributes	attributed to AI; attributed to human; attributed to human+AI	<ul style="list-style-type: none"> <li>“I don't want to, like, downplay undergrad research, but ‘comorbid’ is</li> </ul>

<sup>10</sup> Several existing publications describe the linguistic and stylistic features of AI-generated texts. See, for example, Liang, et al. (2024); Liu, et al. (2024); Rivera Soto, Chen, & Andrews (2025); Smith, et al. (2025).

---

Definition: Qualities that a participant (either orally or textually) attributes to a human and/or nonhuman writer.	something that is usually in higher level research." <ul style="list-style-type: none"> <li>• "I've seen AI work as a bad thesaurus."</li> <li>• "AI is not clever."</li> <li>• "It feels like those are mistakes a person could make."</li> <li>• "I also don't find many undergraduate students who are interested in things like typology - though crazier things have happened."</li> </ul>
Axiological crisis  Definition: When a participant (either orally or textually) expresses some level of discomfort or uncertainty about the triggers, characteristics, and/or attributes they've identified as evidentially significant.	Authorship crisis / doubt / meta-crisis / decision paralysis <ul style="list-style-type: none"> <li>• "The more I try to explain this, the more I think the reverse might be true."</li> <li>• "I want this to have been written by a human because it is more engagingly written, includes a focusing question, mentions actual authors, and does a better job of articulating and forwarding an argument. I fear that it may have been written by AI for all of these reasons."</li> <li>• "I could still be wrong, though."</li> </ul>

---

Inspired by the categories outlined in Table 1, we offer two major themes: First, readers latch onto certain textual cues as triggers that enable them to make assumptions about who/what wrote a text. When combined with prior experiences and knowledge, triggers evidentially warrant decisions that readers make about ambiguously-authored texts. Second, readers implicitly attribute certain characteristics to what they assume is an AI-generator or human author. When readers realize the stakes or consequences of making such characteristics explicit (e.g., revealing value-judgments about certain kinds of writers/writing; being wrong) they may experience what we call an axiological crisis.<sup>11</sup> Next, we explicate the contours of each theme.

**Theme #1. Readers latch onto certain textual cues as triggers that enable them to make assumptions about who/what wrote a text. When combined with prior**

---

<sup>11</sup> See also Hutton and King's (2024) description of our field's "tacit axiologies of reading."

**experiences and knowledge, triggers evidentially warrant readers' decisions about ambiguously-authored texts.**

"There's something fishy going on there."

Synchronous cohort participant

For our purposes, we identified a "trigger moment" when, as the quote above indicates, a participant experienced a kind of "aha!" moment when reading a text; such a moment marked when a participant latched onto one or more "cues in the text" (Haas and Flower, p. 169) that, at least for them, revealed authorship. For example, we coded for triggers when participants pointed to a phrase or word as a "red flag," "signal," or a "hallmark" of either AI's or students' writing style. Other participants reported that a phrase or word "jumped out at" them or inspired a "gut feeling" about authorship. Other times, the trigger moment was more ineffable: "it was odd having, like, I can't really even explain it, but it felt sort of jarring to all of a sudden have this study, kind of in the middle..." The existence of the "trigger" code confirms what Tang, Chuang, and Hu (2024) have argued: AI-generated prose "may leave some distinctive signals in their generated text," which may, indeed, allow readers "to distinguish between LLM [large language models] and human-authored text" (p. 54).

While not all participants explicitly identified a trigger moment (or what Tang, Chuang, and Hu might categorize as a "distinctive signal"), everyone invoked specific (and sometimes overlapping) textual cues that they latched onto as evidentially triggering. Some participants referred to certain words or phrases as sounding "slick," or like "it might have been produced by something that has a very good sense of which words should go together." One participant repeatedly used the word "blip" to describe having been triggered by a cue in the text. When we asked them to describe what they meant by "blip," they said: "a kind of mistake that humans might make naturally, but that a machine is less likely to make." Participants frequently rooted their having been triggered by cues in the text in a previous experience with and/or knowledge about AI:

- "I noticed there are very few grammatical errors in AI generated prose."
- "'Leveraged' is a verb AI uses."
- "I don't see AI using semicolons."

In the above evidential warrants, participants are pointing toward what other researchers might describe as a host of "linguistic patterns," including "vocabulary features, part-of-speech, dependency parsing, sentiment analysis, and stylistic features" (Tang, Chuang, & Hu, p. 54).

Participant responses to "Please explain why/how you categorized Abstract 1, be as specific as possible" and "Please explain why/how you categorized Abstract 2, be as specific as possible" were coded to identify the triggers that participants associated with human-authored vs. AI-generated prose. Then, we took the additional step of dimensionalizing how participants framed specific triggers in terms of valence. For example:

- "+ word use/language" indicates that the participant found the word use/language to be appropriate or effective;

- “~ word use/language” indicates that the participant only noted a particular feature of the vocabulary, but did not comment on its appropriateness or effectiveness; and
- “- word use/language” indicates that the participant found the word use/language to be inappropriate or ineffective.

Figure 2 reports results from a fine-grained analysis of how asynchronous cohort participants framed their triggers in terms of valence—i.e., whether the “cue in the text” was appropriate/effective or inappropriate/ineffective.

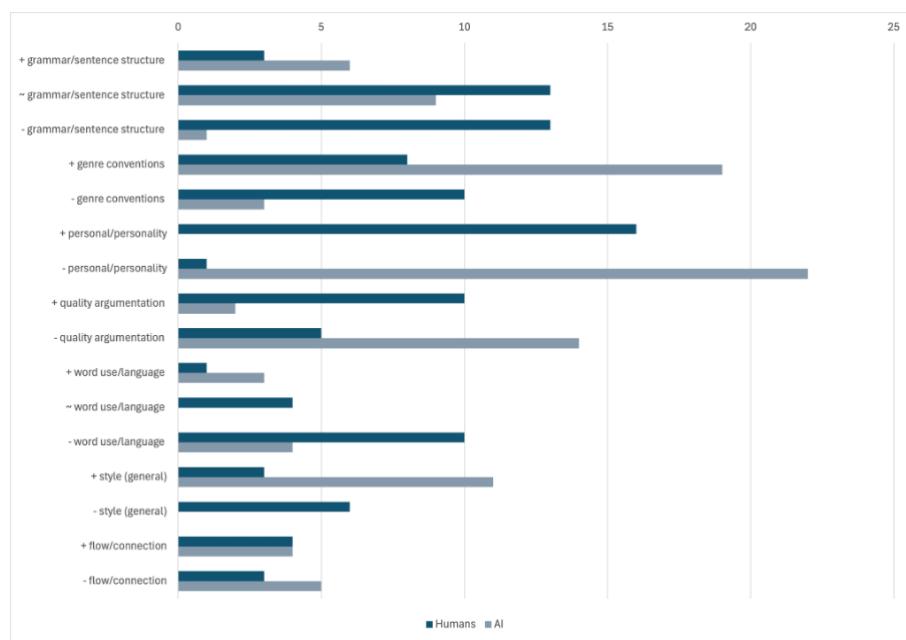


Figure 2. Asynchronous participants' triggers by frequency of application to human writing and AI writing, dimensionalized according to valence

After analyzing both synchronous and asynchronous responses, three triggers, in particular, emerged as most significant, which we unpack in greater detail below. These include:

- “+/- grammar/sentence structure,” and
- “+/- genre conventions”
- “+/- personal/personality”

Our most used trigger, “+/- grammar/sentence structure,” included instances where participants invoked grammar rules to warrant claims about who/what wrote an abstract. Extending what Tang, Chuang, & Hu discovered about the significance of linguistic patterns for detecting AI-generated prose, participants often associated “mechanical errors” with

undergraduate writing and “correct” grammar with AI writing. For example, one participant wrote that “there is one grammatical mistake - a dangling modifier (which I marked as “indicative of human writing”) that I would not expect to see in AI.” This association was so profound that participants used grammatical errors as an AI litmus test. In a synchronous focus group, one participant’s instinct was to start their focus group conversation with the assertion that “I mean, they all seem about equally grammatically right.” When abstracts did not contain errors, participants seemed at a loss for how to proceed with determining authorship.

Other participants latched onto an abstract’s use of “like” rather than ‘such as’ or pointed toward “awkward but not wrong” phrasing to warrant authorship decisions. Another identified an abstract’s reliance on “colons, dashes, and semicolons” and asserted that “students don’t know how to do that.” It is in this trigger where we most clearly see evidence of some readers’ reliance on “deficit discourse” (cf. Shapiro, 2014) concerning student writing (which we briefly discuss in theme #2).

Another trend we noted within this trigger was the assumption that rhetorical questions are perceived as a uniquely human phenomenon. Study stimulus 1 (synchronous cohort) and 2 (asynchronous cohort) both open with a rhetorical question. This was a significant textual cue for participants, with references to it appearing in five out of six focus groups and 11 of 63 survey responses. Universally, participants associated this sentence structure with something “that felt very human to me.” For many, the rhetorical question was determinative: “What really convinced me was the question. AI will write a question if you put it in the prompt, but I feel like an undergraduate would be more likely to include a question in their abstract then [sic] AI.”

The second most common trigger, “+/- genre conventions” represents instances where participants associated either human or AI writing with prose that adhered to or deviated from genre conventions. Interestingly, when dimensionalized for valence, participants did not agree about who/what more frequently failed to follow genre conventions. While some participants recognized that an abstract “follows the CARS model from Swales pretty closely” or adhered to “standard abstract style,” they inconsistently attributed this trigger to undergraduate writers or AI text-generators (Swales and Feak, 2009). In other words, while participants tended to agree that (not) following genre conventions is evidentially meaningful, the evidential *backing* (Toulmin 2003) they supplied when making authorship determinations was contradictory.

Adding greater detail to what Amirjalili, Neysani, and Nikbakht (2024) discovered as it concerns the relationship between authorship and voice, valences for our third most common trigger, “+/- personal/personality” are almost completely segregated. No participants associated AI writing with phrases like “natural” or having a “human voice.”<sup>12</sup> Indeed, participants regularly reflected on how AI writing lacked a certain *je n’ai c’est quoi* that created “a distinct voice” in texts. One participant wrote of Abstract 2 that “It felt like it came from

---

<sup>12</sup> Interestingly, one participant lamented the “robotic” training of student writing in graduate programs.

someones [sic] mind.” On the other side, participants described prose that “has a distance to it, somehow” or that they thought “it sounds soulless.” Perhaps frustratingly, participants did not always specify which parts of the abstract were either missing or abundantly included “personality.” In other words, while some readers warranted authorship assumptions with the personal/personality trigger, the warrant was without evidential backing (Toulmin 2003). Well-warranted claims without Toulminian backing epitomize what perplexes us most about felt sense or tacit knowledge. Do we, indeed, “know more than we can tell” (Polanyi)? Or might we just know less than we think we know?

Within the synchronous cohort, participants relied on terms such as “quirky” when describing human writing, or marked “gobbledygook” as a hallmark of AI-generated text. Wrestling with ambiguity, some focus group participants found themselves asking one another to clarify if they thought the prose was “odd in a human way or odd in an AI way?” “Abstract 2,” one asynchronous participant wrote, “sounded more fluid and natural,” though the same participant later acknowledged that “Abstract 1 also had a few phrases that sounded distinctly human” and concluded with “I’m not sure if I can point to anything specific.” These meanderings perfectly encapsulate, yet again, the way readers’ felt sense resists evidential backing (Toulmin 2003), even if the “personal/personality” trigger serves as a convincing evidential warrant.

No matter which triggers participants evidentially mobilized, they were filtered through the lens of participants’ prior experiences with and/or knowledge about undergraduate writing and/or AI. As Polanyi describes, prior experiences and knowledge are “what we personally bring to the process of understanding or discovery” (qtd. in Turner, p. 185).<sup>13</sup> Occasionally, participants’ knowledge base included popular (mis)conceptions about undergraduate writing or AI text-generation. Often, participants attributed particular terms (e.g., “delve”) or sentence structures (e.g., lists of three; strict parallelism) to AI text-generators based on rumors, articles they’d read, or even AI committee work they’d performed.

Although we’ve chosen not to discuss them at great length, table 2 provides definitions and representative examples for each of the remaining triggers (+/- quality argumentation; +/~/- word use/language; +/- flow/connection; and +/- style [general]).

---

<sup>13</sup> Turner says it this way: “What we see depends on the subsidiaries we possess and which are employed in the course of our meaning-making activities” (p. 186). Similarly, writing studies scholars have referred to “subsidiaries” and prior experiences and knowledge as one’s “repertoire” (McCormick; Haswell, et al.).

Table 2. Four remaining triggers, dimensionalized by valence

Trigger	Definition	Representative Example
+/- quality argumentation	When participants comment on study design, research findings, and/or data-based conclusions	+ "This is a little more even in terms of like study gap methods, findings, implications"  - "I think due to there's no mention of methodology either, which could say, oh, it's poorly written"
+/- word use/language	When participants comment on particular vocabulary choices	+ "Sophistication of vocabulary"  - "The use of the word "utilizing," a jargon term people use when they want to sound more official but which adds no meaning to the sentence"
+/- flow/connection	When participants comment on transitions or organization	+ "but like that...kind of qualification of putting it in parentheses seemed like some kind of, organization of the importance of the qualifier, which I didn't necessarily see the...large language model was offering"  - "The overall flow felt, like, 'Oh, I got distracted by something else'"
+/- style (general)	When participants comment on writing quality in a way that isn't already accounted for by the other triggers	+ "it's a little too good"  - "the sentences are more convoluted and tend to be confusing"

**Theme #2: Readers implicitly attribute certain characteristics to what they assume is an AI-generator or human author. When readers realize the stakes or consequences of making such characteristics explicit (e.g., revealing value-judgments about certain kinds of writers/writing; being wrong) they may experience what we call an axiological crisis.**

In an analytic memo from earlier in our coding process, we wrote:

Perhaps need a code that indexes participants' processing-out-loud of what I'm tentatively calling an "authorship crisis," which basically means the participant is having to reckon with the reality of their implicit assumptions about student vs. AI writing (Participant #8 does this a lot).

Essentially, what we uncovered is that, while engaging with ambiguously-authored prose, readers reasoned about authorship by attributing certain characteristics to what they assume was an AI-generator or human author based on previous experiences and/or felt sense. For example: "When I'm looking at my students' stuff...there's a very, very high level of jargon." Some participants, though, seemed to wrestle with making these attributions out loud: "I don't think AI would make that mistake. Not that it's a mistake, but I don't think it would do that." Upon realizing that they'd attributed "mistakes" to undergraduate student writers, some participants would almost censor or police their assumptions—a phenomenon we initially coded as an authorship- and/or meta-crisis.

As analysts, what we initially termed an authorship- or meta-crisis also occasionally looked like a participant catching themselves, perhaps mid-thought or mid-sentence, after hearing what they just said or wrote, and realizing that their reasoning about authorship may be either tautological and/or rooted in a deficit discourse frame. Some participants would attempt to resolve the axiological crisis by turning toward humor or hedges:

- "I don't...if that...if there's a human in there I don't know what they think is going on....[laughter]"
- "Number three feels like a humanist tried to write an IMRAD abstract."
- "Does this mean this is a human doing a very bad job?"
- "The funny part of this is the more that you work with both of these, the more you could make an argument for any of them."
- "The language in here is maybe a little bit more...'creative'...or what, like, a student might try to do in an abstract."

We developed the axiological crisis category to index when readers realized that their felt sense about certain kinds of writing was actually warranting value-based assumptions that they were not prepared to fully own (out loud). Recognizing the potential stakes in revealing how his authorship assumptions are tethered to what we as writers value, one participant adamantly declared that he was "not at all comfortable making these [authorship] decisions."

Others articulated an axiological crisis by asking hard questions that perhaps reveal an awareness that their authorship hunches were missing evidential backing (Toulmin 2003):

- “But I think the thing that makes this so tricky for me is like, okay, wait, so does that...does that mean this is a...a...a...human or does that mean that this is AI?”
- “Are we assuming the human is the one who did the best job? Are we assuming the human is the one that was more likely to have the errors?”
- “Are you saying ‘odd’ in a human way or an AI way?”
- “‘Adapt and innovate...’ Is that a human touch or a machine?”
- “Is it a sign of a human using strong verbs? Or a large language model...?;
- “I...because...I guess, like, the, the question is like...like...who's the...who's the better writer... and the, you know, who's the smooth? The smoother writer? The skilled human practitioner or the thing?”

After articulating their reasoning, some participants’ axiological crises yielded doubt not just about the reliability of their felt sense, but also about their usefulness to our study:

I am identifying this passage as AI generated only because the other passage sounds much more like an undergraduate. However, it's also possible that an undergraduate would use the words I marked in Abstract 1 in an attempt to sound more serious. If I saw Abstract 1 in isolation, I would be "not at all confident" that it was written by AI. I am also not experienced in reading social science abstracts, so I am probably not so useful for this study!

Still, some participants were less shy about owning the characteristics, attributes, and concomitant value-judgments to which their authorship assumptions were tethered:

I want this to have been written by a human because it is more engagingly written, includes a focusing question, mentions actual authors, and does a better job of articulating and forwarding an argument. I fear that it may have been written by AI for all of these reasons.

Others, however, seemed to tip toe up to an assumption about who or what was more likely to, for example, “use strong verbs,” and then back away before fully explicating what such an attribute might mean about what they, as readers, valued as good writing. Finally, some participants vocalized doubt in the results of their reasoning, altogether: “The more I try to explain this, the more I think the reverse might be true.” In such cases, we wondered whether participants felt themselves reverse-engineering their rationale for why they felt like something was “giving AI.” In other words, the axiological crisis may reveal to a reader the possibility that their authorship assumptions are rooted in tautological reasoning, or at least reasoning that’s missing an important argumentative part of speech: evidential backing (Toulmin 2003).

Axiological crises make us wonder if at least some of the angst around AI in the writing classroom is rooted in the friction produced when readers explicate reasoning about who/what wrote a text. Axiological crises represent quite a conundrum, especially for readers who (think they) should know better. Writing instructors, for example, may feel like they should know better than to make assumptions about student writing that are rooted in deficit discourses. Writing instructors with backgrounds in rhetorical theory may feel like they should know better than to have missing and/or inconsistent evidential backing (Toulmin 2003). When readers are asked to explicate what was once only a "felt sense," they may realize that their explications have consequences. Valence and value-judgments are made manifest through such explications. Consequently, readers may experience guilt, doubt, and/or disappointment in themselves if they detect deficit and/or faulty logic as part of what causes "cues in the text" (Haas and Flower, p. 169) to trigger for them a sense of "offness" (Perl, p. ix).

Even if readers' assumptions about who/what wrote an ambiguously-authored text aren't rooted in dominant discourses around who/what is a "good" writer, expert readers might instead find themselves wrestling with what it might mean if they're *wrong*. The inability to correctly discern who/what has written a text, especially among writing educators, may cause one to feel as though their expertise has waned or been rendered obsolete by AI's presence.

What might it mean, however, to see the axiological crisis as an opportunity for growth? Our study allowed participants not just to language-out-loud their felt sense about an ambiguously-authored text; it also provided readers with an opportunity to metacognitively reflect on their reasoning. Adler-Kassner, Majweski, and Koshnik (2012) note that as "knowledge...becomes less tacit and more explicit, discursive, and conscious, at least for a time," readers "not only know what they know, but they are also more likely to recognize how they know it" (n.p.). Might there be value, then, in collaboratively reading ambiguously-authored texts and discussing how and why certain textual cues trigger our literate sensibilities? Even if it doesn't make us more expert in "detecting" AI-generated prose, perhaps reading ambiguously-authored texts provides readers with an opportunity to grapple with their own unique axiological crises. Turner, channeling Polanyi, goes so far as to say that "the process of scientific discovery itself may be understood as a process in which our tacit knowledge, for example the unarticulated intuitions that lead to a discovery, is made explicit" (Turner, p. 183).

#### 4. Next Steps and Limitations

Even while acknowledging that it was difficult, many participants ended the study by expressing enthusiasm about the reading activity we invited them to complete. Several especially appreciated the opportunity to process their felt senses collaboratively. Some even said they had fun! Without these low-stakes opportunities for working one's way through an axiological crisis, we worry that more educators will proudly declare, as one participant did, that they "have gotten \*something\* out of filing all those academic misconduct reports!" Perhaps punitive attitudes toward students who may use AI text-generators could be

tempered by offering instructors an opportunity to metacognitively process the felt senses that evidentially warrant their authorship assumptions.

We offer this recommendation tentatively because our study has numerous limitations. First, the relatively small sample size ( $n=76$ ) prevents us from making statistical inferences. This was not an experimental study with official treatment and control groups. We also chose not to include so-called AI detectors as part of our study design since we're focused exclusively on *humans'* reading processes. As such, we're limited to building only substantive, local theories.

A larger, more diverse participant pool is certain to yield different results. Notably, our participant pool was limited to those working and/or studying within higher education. If nothing else, future studies should incorporate a wider array of disciplines whose class sizes and labor conditions may offer greater nuance. Future studies might also consider how lay publics read ambiguously-authored texts. Employing different study stimuli (e.g., multimodal texts, such as infographics and public service announcements, etc.) would also add another dimension.

And, of course, our methods have consequences for what we can argue. Focus groups allowed us to capture rich discussions between participants, but not all participants were comfortable with reasoning out loud in front of others. Some participants, in fact, expressed that they had difficulty concentrating on the reading task in a room with other participants. Everything we know about literate practice points to how important material conditions are; participants likely would have articulated different triggers had they been able to read in their preferred environment.

Similarly, our study's survey instruments, while allowing for broader and more convenient data collection, limited the depth of information we were able to capture. Unlike our synchronous cohort, where we included space for semi-structured interviews, asynchronous participants were limited to text boxes and unable to ask us questions. We also wonder if our study population was perhaps predisposed toward a kind of cynicism about the efficacy of AI text detection software and so-called committees of academic misconduct who mediate policy-violations concerning students' use of AI. We found ourselves wondering how our findings may have been different had our participants hailed from, for example, our institution's College of Pharmacy, where instructors are reading essays on the same topic from class sizes of 100+ students (rather than various essays from undergraduate English classes that are capped at 20-24 students). Might the axiological crises have looked different among such participants? Would there have been an axiological crisis at all? As an exploratory study, though, we learned a lot—not just about how readers make decisions about ambiguously-authored texts, but also about how rich readers' reasoning processes are, especially when we account for felt sense and some readers' eventual recognition that the evidential backing they explicate in support of authorship claims may be missing and/or inconsistent.

We've begun preliminary analyses of video data from the synchronous cohort and plan to code exclusively for gesture and other forms of nonverbal communication. Our future findings should, therefore, help to deepen understanding of the role of embodied knowledge when

readers collaboratively negotiate about who and/or what wrote a text. We're also eager to use our RX Lab's new Tobii Eye Tracking station to capture an even more fine-grained sense of how readers' corporeal bodies affect literate practices (cf. Anson & Schwegler). Studies designed to account for transnational and/or multilingual readers' practices would also help to paint a fuller picture of what's required when discerning human- from AI-generated prose.

## References

Adelani, D. I., Mai, H., Fang, F., Nguyen, H. H., Yamagishi, J., & Echizen, I. (2020). Generating sentiment-preserving fake online reviews using neural language models and their human-and machine-based detection. In *Advanced information networking and applications: Proceedings of the 34th international conference on advanced information networking and applications* (AINA-2020) (pp. 1341-1354). Springer International Publishing. [https://doi.org/10.1007/978-3-030-44041-1\\_114](https://doi.org/10.1007/978-3-030-44041-1_114)

Adler-Kassner, L., Majewski, J., & Koshnick, D. (2012). The Value of Troublesome Knowledge: Transfer and Threshold Concepts in Writing and History. *Composition Forum* (Vol. 26). Association of Teachers of Advanced Composition.

Adler-Kassner, L., Clark, I., Robertson, L., Taczik, K., & Yancey, K. B. (2016). Assembling knowledge: The role of threshold concepts in facilitating transfer. *Critical transitions: Writing and the question of transfer*, 17-47. <https://doi.org/10.37514/PER-B.2016.0797.2.01>

Amidon, T. R. (2022, June). Troubling the Tacit: A Review Essay of Harry Collins's (2010) *Tacit and Explicit Knowledge*. In *Composition Forum* (Vol. 49).

Amirjalili, F., Neysani, M., & Nikbakht, A. (2024). Exploring the boundaries of authorship: A comparative analysis of AI-generated text and human academic writing in English literature. *Frontiers in Education*, 9. <https://doi.org/10.3389/feduc.2024.1347421>

Anson, C. M., & Schwegler, R.A. (2012). Tracking the Mind's Eye: A New Technology for Researching Twenty-First-Century Writing and Reading Processes. *College Composition and Communication* 64(1), 151-171. <https://doi.org/10.58680/ccc201220864>

Bedington, A., Halcomb, E., McKee, H. A., Sargent, T., & Smith, A. (2024). Writing with generative AI and human-machine teaming: Insights and recommendations from faculty and students. *Computers and Composition*, 71. <https://doi.org/10.1016/j.compcom.2024.102827>

BBC. 13 TikTok Phrases You Need to Know. <https://www.bbc.co.uk/bitesize/articles/zrft9ty#zxj9p9q>

Burriss, S. K., Smith, B. E., Shimizu, A. Y., Hundley, M., Pendergrass, E., & Molvig, O. (2025). Exploring the ethics of multimodal composition with AI: Student and educator perspectives on evaluating and using generative models in the classroom. *Contemporary Issues in Technology and Teacher Education*, 25(2).

Carillo, E. C. (2015). *Securing a place for reading in composition: The importance of teaching for transfer*. University Press of Colorado. <https://doi.org/10.7330/9780874219609>

Clayson, A. (2018). Distributed cognition and embodiment in text planning: A situated study of collaborative writing in the workplace. *Written Communication*, 35(2), 155-181. <https://doi.org/10.1177/0741088317753348>

Collins, H. (2019). *Tacit and explicit knowledge*. University of Chicago press.

Crozier, M., & Workman, E. (2022). Discourse-based interviews in institutional ethnography: Uncovering the tacit knowledge of peer tutors in the writing center. In *Composition Forum* (Vol. 49). Association of Teachers of Advanced Composition.

Cummings, R. E., Monroe, S. M., & Watkins, M. (2024). Generative AI in first-year writing: An early analysis of affordances, limitations, and a framework for the future. *Computers and Composition*, 71. <https://doi.org/10.1016/j.compcom.2024.102827>

Dugan, L., Ippolito, D., Kirubarajan, A., & Callison-Burch, C. (2020). RoFT: A tool for evaluating human detection of machine-generated text. <https://doi.org/10.18653/v1/2020.emnlp-demos.25>

Evia, C., & Patriarca, A. (2012). Beyond compliance participatory translation of safety communication for Latino construction workers. *Journal of Business and Technical Communication*, 26(3), 340–367. <https://doi.org/10.1177/1050651912439697>

Fleckenstein, K. S. (2003). *Embodying literacies: Imageword and a poetics of teaching*. SIU press.

Glaser, B., & Strauss, A. (1967). *The Discovery of Grounded Theory: Strategies for Qualitative Research*. Sociology Press. <https://doi.org/10.1097/00006199-196807000-00014>

Haas, C., & Flower, L. (1988). Rhetorical Reading Strategies and the Construction of Meaning. *College Composition and Communication*, 39(2), 167-183. <https://doi.org/10.1177/105065190101500402>

Haas, C., & Witte, S. P. (2001). Writing as an embodied practice: The case of engineering standards. *Journal of Business and Technical Communication*, 15(4), 413-457.

Haswell, R. H., Briggs, T. L., Fay, J. A., Gillen, N. K., Harrill, R., Shupala, A. M., & Trevino, S. S. (1999). Context and rhetorical reading strategies: Haas and Flower (1988) revisited. *Written Communication*, 16(1), 3-27. <https://doi.org/10.1177/0741088399016001001>

Hayles, N.K. (1999). *How We Became Posthuman. Virtual Bodies in Cybernetics, Literature, and Informatics*. The University of Chicago Press. <https://doi.org/10.7208/chicago/9780226321394.001.0001>

Herring, K. D. (2024). It's giving queer rhetorical pedagogy: Introducing rhetorical criticism with queer vernacular. *Communication Teacher*, 38(3), 260-267. <https://doi.org/10.1080/17404622.2024.2342803>

Hutton, L., & King, C. M. (2024). A Commonplace Problem: Uncovering Composition's Tacit Axiologies of Reading. *College Composition & Communication*, 76(1), 90-119. <https://doi.org/10.58680/ccc202476190>

Keller, D. (2013). *Chasing literacy: Reading and writing in an age of acceleration*. University Press of Colorado. <https://doi.org/10.7330/9780874219333>

Kim, S., & Cho, S. (2017). How a tutor uses gesture for scaffolding: A case study on L2 tutee's writing. *Discourse Processes*, 54(2), 105–123. <https://doi.org/10.1080/0163853X.2015.1100909>

Knowles, A. M. (2024). Machine-in-the-loop writing: Optimizing the rhetorical load. *Computers and Composition*, 71. <https://doi.org/10.1016/j.compcom.2024.102826>

Laquintano, T. P., Schnitzler, C., & Vee, A. (Eds.). (2023). "An Introduction to Teaching with Text Generation Technologies" In *TextGenEd: Teaching with Text Generation Technologies*. WAC Clearinghouse.

Lee, M., Liang, P., & Yang, Q. (2022, April). Coauthor: Designing a human-ai collaborative writing dataset for exploring language model capabilities. In *Proceedings of the 2022 CHI conference on human factors in computing systems* (pp. 1-19). <https://doi.org/10.1145/3491102.3502030>

Liang, W., Izzo, Z., Zhang, Y., Lepp, H., Cao, H., Zhao, X., ... & Zou, J. Y. (2024). "Monitoring ai-modified content at scale: A case study on the impact of chatgpt on ai conference peer reviews." *Proceedings of the 41st International Conference on Machine Learning*, Vienna, Austria. arXiv:2403.07183.

Liu, J. Q. J., Hui, K. T. K., Al Zoubi, F., Zhou, Z. Z. X., Samartzis, D., Yu, C. C. H., Chang, J. R., & Wong, A. Y. L. (2024). The great detectives: Humans versus AI detectors in catching large language model-generated medical writing. *International Journal for Educational Integrity*, 20(1). <https://doi.org/10.1007/s40979-024-00155-6>

McCormick, K. (1989). Expanding the Repertoire: An Anthology of Practical Approaches for the Teaching of Writing (Reading-to-Write Report No. 11). Technical Report No. 30.

Odell, L., Goswami, D., & Herrington, A. (1983). The discourse-based interview: A procedure for exploring the tacit knowledge of writers in nonacademic settings. In P. Mosenthal, L. Tamor, & S. A. Walmsley (Eds.), *Research on writing: Principles and methods* (pp. 221–236). Longman.

Olinger, A. R. (2014). On the instability of disciplinary style: Common and conflicting metaphors and practices in text, talk, and gesture. *Research in the Teaching of English*, 48(4), 453-478. <https://doi.org/10.58680/rte201425162>

Pandey, H. L., Bhusal, P. C., & Niraula, S. (2024). Large language models and digital multimodal composition in the first-year composition classrooms: An encroachment and/or enhancement dilemma. *Computers and Composition*, 75. <https://doi.org/10.1016/j.compcom.2024.102892>

Perl, S. (1979). The composing processes of unskilled college writers. *Research in the Teaching of English*, 13(4), 317-336. <https://doi.org/10.58680/rte201117867>

Perl, S. (1980). Understanding composing. *College Composition & Communication*, 31(4), 363-369. <https://doi.org/10.58680/ccc198015928>

Perl, S. (2004). *Felt sense: Writing with the body*. Portsmouth, NH: Boynton/Cook Heinemann.

Polanyi, M. (1958). *Personal knowledge*. Routledge.

Rice, J. (2015). Para-expertise, tacit knowledge, and writing problems. *College English*, 78(2), 117-138. <https://doi.org/10.58680/ce201527547>

Rivera Soto, R., Chen, B., & Andrews, N. (2025). Language models optimized to fool detectors still have a distinct style (and how to change it). arXiv. <https://doi.org/10.48550/arXiv.2505.14608>.

Rule, H. J. (2017). Sensing the sentence: An embodied simulation approach to rhetorical grammar. *Composition Studies*, 45(1), 19-38.

Saldaña, J. (2009). *The coding manual for qualitative researchers*. Sage.

Sauer, B. A. (2003). *The rhetoric of risk: Technical documentation in hazardous environments*. Routledge. <https://doi.org/10.4324/9781410606815>

Shapiro, S. (2014). "Words that you said got bigger": English language learners' lived experiences of deficit discourse. *Research in the Teaching of English*, 48(4), 386-406. <https://doi.org/10.58680/rte201425159>

Smith, B., Bouadjenek, M. R., Kheya, T. A., Dawson, P., & Aryal, S. (2025). A comprehensive analysis of large language model outputs: Similarity, diversity, and bias. arXiv. <https://doi.org/10.48550/arXiv.2505.09056>

Swales, J. M., & Feak, C. B. (2010). From text to task: Putting research on abstracts to work. In *English for professional and academic purposes* (pp. 169-182). Brill. [https://doi.org/10.1163/9789042029569\\_012](https://doi.org/10.1163/9789042029569_012)

Tang, R., Chuang, Y. N., & Hu, X. (2024). The science of detecting LLM-generated text. *Communications of the ACM*, 67(4), 50-59. <https://doi.org/10.1145/3624725>

Toulmin, S. E. (2003). *The uses of argument*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511840005>

Turner, S. (2022). Polanyi and tacit knowledge. In *The Routledge Handbook of Philosophy and Implicit Cognition* (pp. 182-190). Routledge. <https://doi.org/10.4324/9781003014584-17>

Valiaiev, D. (2024). Detection of machine-generated text: Literature survey. arXiv preprint. arXiv:2402.01642.

Vee, T. Laquintano, & C. Schnitzler (2023), *TextGenEd: Teaching with Text Generation Technologies*. The WAC Clearinghouse. <https://doi.org/10.37514/TWR-J.2023.1.1.02>

Walters, W. H. (2023). The effectiveness of software designed to detect AI-generated writing: A comparison of 16 AI text detectors. *Open Information Science*, 7(1), 20220158. <https://doi.org/10.1515/opis-2022-0158>

Wolfe, J. (2005). Gesture and collaborative planning: A case study of a student writing group. *Written Communication*, 22(3), 298-332. <https://doi.org/10.1177/0741088305278108>

## Appendix A: Synchronous Study Stimuli

### Stimulus 1: Human-Authored

#### Source:

Drewry, D. J. and Z. Reese (2021). "The Effect of Perceived Uncertainty on Competitive Behavior." *American Journal of Undergraduate Research* 18(3). <https://doi.org/10.33697/ajur.2021.049>

How do people behave in the face of uncertainty? Some studies suggest that even when they are unaware of how others will behave, people default to cooperative behavior; however, other research suggests that uncertainty leads to more competitive behavior. Little research has examined how individual differences moderate such behavioral decisions. This study proposes that a stable (dispositional) sense of justice may, ironically, lead to more competitive behavior. Specifically, people who score highly in belief in a just world, system justification, and religiosity, and low in ambiguity tolerance may be more inclined to compete rather than cooperate because they believe people who experience positive outcomes deserve those outcomes regardless of the means taken to achieve them. Across two studies, participants (N = 288) engaged in a prisoner's dilemma game — a task where they must choose to compete or cooperate — and completed the aforementioned individual difference measures. Results show that people tended to cooperate, but those high in system justification and belief in a just world were more likely to compete. In other words, people with a strong sense of cosmic justice were likely to exhibit competitive behavior under uncertain conditions.

### Stimulus 2: AI + Human Co-Written

#### Source:

Graham A., D. R. Terrill, S. I. Boyd, I. Benjamin, M. Ward, and M. Zimmerman, (2021). "Treatment Outcomes in a Partial Hospital Program for Patients with Social Anxiety Disorder: The Effects of Comorbid Major Depression." *American Journal of Undergraduate Research* 17(4). <https://doi.org/10.33697/ajur.2020.033>

This study explores the impact of comorbid major depressive disorder (MDD) on treatment outcomes in patients with social anxiety disorder (SAD) in a partial hospitalization setting, addressing inconsistencies in previous findings. Prior research has shown that MDD can hinder treatment efficacy for SAD, but much of this work has focused on inpatient and outpatient contexts, yielding mixed results. This study compared pre- and post- treatment measures of anxiety and depression in two groups of patients at Rhode Island Hospital: those with only SAD and those with both SAD and MDD. Results revealed that patients with comorbid SAD and MDD experienced significantly less improvement in anxiety symptoms and constructs related to depression remission, though their progress on core depression symptoms did not differ significantly from those with SAD alone. These findings highlight the nuanced impact of

MDD on SAD treatment outcomes and suggest that clinicians should consider comorbidity effects when designing and implementing treatment plans for this population.

*Prompt (ChatGPT4, generated 10/31/24): Revise the following abstract. The final abstract should be between 170-200 words [pasted text of the original abstract].*

*Original Abstract:* Although previous studies have examined the impact of comorbid major depressive disorder (MDD) on social anxiety disorder (SAD), the results have been somewhat mixed. Furthermore, most studies have been conducted in inpatient or outpatient settings. Given the large body of literature that suggests that this particular comorbidity can have negative effects on treatment efficacy and outcomes, it is important to continue to explore its impact. The present study aims to clarify contradictions in the literature and expands on previous studies by examining patients in a partial hospitalization setting. Patients at Rhode Island Hospital with a diagnosis of SAD were compared to those with comorbid SAD and MDD on pre-treatment and post-treatment measures of anxiety and depression. The results indicated that while the comorbid group showed significantly less improvement post-treatment on anxiety symptoms and constructs related to remission from depression, they did not show significantly less improvement on depression symptoms. The implications of these results for clinical practice are discussed.

### **Stimulus 3: AI-Written (ChatGPT4)**

#### **Source:**

Gilliam, H. (2020). "Travel Through Time: From 9/11 to COVID-19, Parallel Predictive Analysis of Travel Marketing." *American Journal of Undergraduate Research* 17(3). <https://doi.org/10.33697/ajur.2020.029>

This research examines the travel industry's response to two crises, 9/11 and COVID-19, both of which induced significant changes in travel marketing. After 9/11, heightened security protocols and public fear altered the travel landscape. Companies had to reassure consumers about safety while promoting travel within a framework of national unity. Similarly, COVID-19 imposed a global public health crisis, leading to widespread travel restrictions, economic disruptions, and shifts in consumer behavior. By comparing the responses to these events, the study identifies parallel marketing strategies that leveraged themes like community resilience, safety, and support for local tourism. This paper suggests that post-COVID-19, travel marketing will echo strategies used after 9/11, focusing on themes of safety confidence, localism, and the American Dream. Messaging will likely promote reconnecting with loved ones, exploring familiar locales in new ways, and supporting local businesses. These insights aim to provide a predictive analysis for travel industry stakeholders on effectively re-engaging consumers in a post-pandemic world, with messaging that resonates with heightened emotional and psychological considerations around safety and community.

*Prompt (ChatGPT4, generated 10/31/24): Generate a 175-200 word abstract based on this paper. [Upload entire paper]*

*Original Abstract:* The events of 9/11 drastically changed the state of the nation across many industry sectors, with the tourism industry among those most affected. Following that horrific day, the nation experienced heightened security measures and protocol, such that the travel industry and travelers would never look the same. People were fearful and anxious, and the tourism industry had to take quick, effective measures to evaluate the consumer response, set a marketing strategy, and promote within a changed national ethos and expectations. COVID-19 is a similar catastrophic, global, and long-term crisis that set our nation on a similarly drastic change in practice and protocol; fear and anxiety were higher than ever. COVID-19 and 9/11 are highly comparable in their market response. By comparing the two events and analyzing the consumer response and advertising messaging, specifically during the stay at home order, a theme and direction for messaging within the travel industry post-COVID-19 can be predicted based on the culture and spirit of The American Dream, confidence in safety, we are in this together, support local tourism, explore your city in a new way, and connect with those you missed.

#### **Stimulus 4: AI-Written (Gemini)**

##### **Source:**

Gilliam, H. (2020). "Travel Through Time: From 9/11 to COVID-19, Parallel Predictive Analysis of Travel Marketing." *American Journal of Undergraduate Research* 17(3). <https://doi.org/10.33697/ajur.2020.029>

The 9/11 terrorist attacks and the COVID-19 pandemic represent two significant events that have profoundly impacted the global tourism industry. Both crises triggered widespread fear, uncertainty, and travel restrictions, forcing the industry to adapt and innovate to survive. By examining the consumer response and advertising messaging during these periods, we can identify key themes and strategies that have proven effective in rebuilding trust and stimulating travel demand. This research delves into the parallels between the two events, focusing on the consumer sentiment, industry response, and advertising campaigns. We analyze how the travel industry leveraged the American Dream, emphasized safety and security, fostered a sense of community, promoted local tourism, and encouraged exploration of familiar destinations in new ways. By understanding these strategies, we can gain valuable insights into the future of the tourism industry and develop effective marketing approaches to revitalize travel in the post-COVID-19 era.

*Prompt (GoogleGemini, generated 10/31/24): Based on the following summary, write 175 word abstract for this paper. [Upload original abstract]*

*Original Abstract:* The events of 9/11 drastically changed the state of the nation across many industry sectors, with the tourism industry among those most affected. Following that horrific day, the nation experienced heightened security measures and protocol, such that the travel industry and travelers would never look the same. People were fearful and anxious, and the tourism industry had to take quick, effective measures to evaluate the consumer response, set a marketing strategy, and promote within a changed national ethos and expectations. COVID-19 is a similar catastrophic, global, and long-term crisis that set our nation on a similarly drastic change in practice and protocol; fear and anxiety were higher than ever. COVID-19 and 9/11 are highly comparable in their market response. By comparing the two events and analyzing the consumer response and advertising messaging, specifically during the stay at home order, a theme and direction for messaging within the travel industry post-COVID-19 can be predicted based on the culture and spirit of The American Dream, confidence in safety, we are in this together, support local tourism, explore your city in a new way, and connect with those you missed.

**Appendix B: Synchronous Cohort: Screening Survey**

Q1 Name

Q2 Email address - if you are selected for participation in this study, we will contact you at the provided email address

Q2 Select your age group

- 18-25
- 26-35
- 36-45
- 46-55
- 56-65
- 65+

Q3 Indicate your gender identity

- Man
- Woman
- Gender Nonconforming/Nonbinary
- Prefer to self-describe [describe here]
- Prefer not to disclose

Q4 Please select the role that best describes you

- Tenure-line Faculty (Assistant, Associate, Full)
- Lecturer, Clinical, or Associated Faculty
- Graduate Student (with one full semester of teaching experience)
- Graduate Student (without one full semester of teaching experience)
- Undergraduate Student

Q9 Please indicate your experience level with generative AI tools—including ChatGPT, Google Gemini, and/or Microsoft Copilot.

▼ I do not know what these tools are ... Use on a regular basis and/or as a part of my regular work

Q10 Please indicate your confidence level in identifying texts that have been authored by a generative AI tool (such as ChatGPT, Google Gemini, Microsoft Copilot, etc).

▼ I do not know what these tools are ... Very Confident

Q11 Please indicate your experience level in writing texts with the assistance of a generative AI tool (such as ChatGPT, Google Gemini, Microsoft Copilot, etc).

▼ I do not know what these tools are ... Very Experienced

Q12 Please indicate your experience level in reading abstracts from social scientific articles.

▼ Not Experienced ... Very Experienced

Q13 Please indicate your experience level in writing abstracts for social scientific articles.

▼ Not Experienced ... Very Experienced

### **Appendix C: Synchronous Cohort: Focus Group Protocol**

I'm going to start by asking each of you a few questions about your position here at OSU. I will start with [participant]. You previously indicated that you are (non-teaching graduate student/teaching graduate student/lecturer/faculty member), correct?

- What classes have you previously taught at OSU?/What classes have you taught in the last 2-3 years?
- (SKIP for faculty) Have you taught classes at any other universities? If yes, do you remember which ones?
- Do you currently have a policy regarding student-use of AI in your classroom?
- Do you currently use any AI detection software like TurnItIn or Honorlock?

Today, we will ask you to read three abstracts from three social scientific undergraduate research papers. You do not need to be familiar with the specific subject matter for this study. One abstract was authored by a person, one was generated by an AI writing tool, and one is by a combination of human authorship and AI generation.

Once I leave the room, I am going to ask you to start reading the text. After you have read all three texts, you will be asked to choose which text is authored by a person, which was generated by an AI tool, and which is a combination of the two. You will be expected to work together to make this determination.

During this time, we will be observing you from a different room. Your comments and interactions will be recorded.

After you read through the text, we will come back into the room to interview each of you individually about your experience. As before, your responses will be recorded.

Do you have any questions about what you are being asked to do?

Is there anything you want to share with us before beginning?

***[Hand Over Selected Stimuli]***

You may begin reading once I leave the room. Once you finish reading, you should state that you are done.

#### **Appendix D: Synchronous Cohort: Semi-Structured Interview Questions**

##### General Questions:

1. Do you believe that this text was (human-authored/AI-authored/or authored by some combination of the two)?
2. Why do you believe that the text was (human-authored/AI-authored/combination-authored)?
3. How did you decide whether this text was (human-authored/AI-authored/combination-authored)?
4. Were there any particular words, sentences, or other textual cues that helped you discern whether this text was (human-authored/AI-authored/combination-authored)?
5. Which is most credible/most knowledgeable/most professional?

##### [Determined AI] Specific Questions:

6. (If Instructor) How would you respond to a student who turned this text [choose one] in for a grade?
7. (If Student) How would you respond to a peer who turned this text [choose one] in for peer review?
8. What recommendations for revising this text [choose one] would you share with the author of this text [choose one]?

##### Conclusion

9. 1 - 10 scale, how confident are you on your conclusion?
10. Is there anything else you’d like to share with us today about your experiences when reading these texts?

**Appendix E: Asynchronous Cohort Survey**

Q6 Select your age group

- 18-25
- 26-35
- 36-45
- 46-55
- 56-65
- 65+

Q7 Indicate your gender identity

- Man
- Woman
- Gender Nonconforming/Nonbinary
- Prefer to self-describe [describe here]
- Prefer not to say

Q8 Optional: If you are willing to share this information, please write the name of your affiliated university *Note: By doing so, your identifying information will be tied to your survey results.*

Q9 Please select the role that best describes you

- Tenure-line Faculty (Assistant, Associate, Full)
- Lecturer, Clinical, or Associated Faculty
- Graduate Student (with one full semester of teaching experience)
- Graduate Student (without one full semester of teaching experience)
- Undergraduate Student

End of Block: Demographic Information

Start of Block: Regarding experience with generative AI tools

Q10 Please indicate your experience level with generative AI tools—including ChatGPT, Google Gemini, and/or Microsoft Copilot.

▼ I do not know what these tools are ... Use on a regular basis and/or as a part of my regular work

Q11 Please indicate your confidence level in identifying texts that have been authored by a generative AI tool (such as ChatGPT, Google Gemini, Microsoft Copilot, etc).

▼ I do not know what these tools are ... Very Confident

Q12 Please indicate your experience level in writing texts with the assistance of a generative AI tool (such as ChatGPT, Google Gemini, Microsoft Copilot, etc).

▼ I do not know what these tools are ... Very Experienced

Q13 Please indicate your experience level in reading abstracts from social scientific articles.

▼ Not Experienced ... Very Experienced

Q14 Please indicate your experience level in writing abstracts for social scientific articles.

▼ Not Experienced ... Very Experienced

Q15 We will now ask you to read/review two undergraduate research abstracts. You do not need to be familiar with the particular subject matter to participate in this survey. One abstract was written by an undergraduate student and one was generated by an AI writing tool. After reading through both abstracts, please indicate which you believe was written by an undergraduate student and which you believe was generated by an AI writing tool. As you read through the abstracts, you can use the highlight tool to select words or sentences you believe are "indicative of AI generation," "indicative of human writing," or "notable for some other reason." We will also ask you some specific questions about your experience.

Q18 How confident are you in that choice?

- not at all confident
- slightly confident
- somewhat confident
- fairly confident
- completely confident

Q19 Please explain why/how you categorized Abstract 1, be as specific as possible

---

---

---

---

---

Q25 If you labelled any text in Abstract 1 as "notable for other reason," please explain

---

---

---

---

Q20 Please indicate whether or not you agree with the following statements about Abstract 1

	Strongly disagree	Somewhat disagree	Neither agree nor disagree	Somewhat agree	Strongly agree
I find the author to be credible	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I find this author to be knowledgeable about the subject	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I find this abstract to be professionally written	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Q21 I think that Abstract 2 was written/generated by

- An Undergraduate Student
- An AI Text Generation Tool

Q23 How confident are you in that choice?

- not at all confident
- slightly confident
- somewhat confident
- fairly confident
- completely confident

Q24 Please explain why/how you categorized Abstract 2, be as specific as possible

---



---



---



---

Q26 If you labelled any text in Abstract 2 as "notable for other reason," please explain

---



---



---



---

Q22 Please indicate whether or not you agree with the following statements about Abstract 2

	Strongly disagree	Somewhat disagree	Neither agree nor disagree	Somewhat agree	Strongly agree
I find the author to be credible	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I find this author to be knowledgeable about the subject	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I find this abstract to be professionally written	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

#### Appendix F: Asynchronous Cohort Stimuli

##### Source:

Jerald, A. (2020). "Exploring the Relationship between Dystopian Literature and the Activism of Generation Z Young Adults." *American Journal of Undergraduate Research* 16(4). <https://doi.org/10.33697/ajur.2020.009>.

##### Original Abstract:

Some recent research has posited that the independent and revolutionary traits of Generation Z can be traced to the circumstances of their births, specifically the 9/11 attacks and the Great Recession. While there has been research examining the effect of these events on the type of behavior Generation Z exhibits towards political and societal issues, there has been little research that examines the literary culture in which they grew up. Did popular dystopian works such as *Catching Fire* by Suzanne Collins (2009), *Divergent* by Veronica Roth (2011), and *The Maze Runner* by James Dashner (2009) have an impact on their political identities and behaviors? This paper examines that question by using a mixed method approach: a public questionnaire, thirteen in-depth interviews with a select group of Generation Z students from the University of Georgia, and direct content analyses of the key works under consideration. This study argues that the relationship between dystopian literature and young adult activism may offer insight into the ways literature can be used as a revolutionary tool. This study also hopes to add to the literature exploring the characteristics of Generation Z and the significance dystopian literature may have not only on a young adult's thoughts but also their actions.

**ChatGPT:**

This study examines the influence of dystopian literature on the activism of Generation Z young adults, who have been shaped by a unique historical context marked by events like the 9/11 attacks and the Great Recession. With limited research on how literature may impact political behaviors, this paper explores whether popular dystopian works, such as *Catching Fire*, *Divergent*, and *The Maze Runner*, contribute to the social and political awareness of Generation Z. Utilizing a mixed-methods approach, the study includes a public questionnaire, in-depth interviews with thirteen Generation Z students from the University of Georgia, and content analysis of the selected novels. The findings suggest that dystopian literature resonates with Generation Z by reflecting contemporary societal issues and fostering empathy, self-reflection, and activism. The research proposes an activist typology within Generation Z influenced by themes in dystopian literature: the selfless activist, the reluctant activist, and the instinctual activist. These typologies reflect how literature may serve as a revolutionary tool, encouraging young readers to challenge societal norms and advocate for change, providing insight into the broader social impact of the humanities.

*Prompt (ChatGPT4, generated 10/31/24): Generate an abstract based on this paper. The abstract should be between 175-200 words. [Upload entire paper]*