

# Ethical Boundaries in the Digital Age: EFL Higher Education Students' Experiences with AI-Assisted Writing

Haitham Ali Althubaiti<sup>1\*</sup>

<sup>1</sup> Assistant Professor of English Language Curriculum and Instruction, Department of Curriculum and Instruction, College of Education, University of Jeddah, Jeddah, Saudi Arabia, Email: Haalthubaiti@uj.edu.sa, ORCID: <https://orcid.org/0000-0001-7755-1220>

## HOW TO CITE:

Haitham Ali Althubaiti (2026). Ethical Boundaries in the Digital Age: EFL Higher Education Students' Experiences with AI-Assisted Writing International Journal of Special Education, 41(19s), 1494 - 1509

## ABSTRACT:

Generative artificial intelligence (AI) has transformed academic writing for English as a foreign language (EFL) learners, yet there is a scarcity of qualitative research regarding their perceptions of the ethical boundaries of AI usage in academic writing. This qualitative research is based on Rest's (1986) Four-Component Model of ethical decision-making, which examines EFL higher education students' interpretation, negotiation, and experiential encounters with ethical AI-assisted writing. Data was gathered via semi-structured interviews with six Saudi EFL master's students and analysed thematically employing a hybrid deductive-inductive methodology. Results suggest that participants see "ethics" as "appropriate application," define ethical usage as support rather than substitution, and adjust acceptable dependence based on task type and significance. The employment of AI is motivated by language uncertainty, time constraints, and routine, and has been normalised to the extent that it circumvents intentional decision-making. Significantly, the lack of explicit institutional guidelines, individuals align their integrity measures with detection technology, with one individual deliberately diminishing her English proficiency to prevent genuine writing from being mistakenly categorised as AI-generated. The research argues that detection tools misclassify EFL students' writing, making enforcement by detection an unreliable basis for academic integrity policy. It proposes that instructors should instead set explicit rules for their own courses, distinguishing acceptable from unacceptable uses of a tool by what the student uses it to do. The ramifications for theoretical frameworks, policy formulation, and educational practices are examined.

**Keywords:** generative artificial intelligence, scholarly honesty, English as a Foreign Language, moral judgement, ChatGPT, artificial intelligence identification.

## COPYRIGHT STATEMENT:

Copyright: © 2026 Authors.

Open access publication under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).

## Introduction

The public release of ChatGPT in late 2022, along with the accelerated deployment of new generative AI tools will likely reshape the way in

which students create academic writing at universities. For students learning English as a Foreign Language (EFL), these tools — including ChatGPT, Grammarly, and QuillBot — can help

with issues such as word choice, register, style, wording, and sentence structure/form – all areas that present major challenges for EFL writers. Thus, writing among EFL learners is moving toward a model that combines human-written content with AI technologies, creating collaborative forms of writing (Barrot, 2023; Kohnke et al., 2023); however, institutions continue to struggle in defining whether specific uses of AI will meet their expectations of academic integrity.

The combination of the above two trends presents EFL writers with a double burden. First, they must deal with the linguistic requirements of writing in Academic English. Second, they must find ways to determine what constitutes permissible AI-assistance use within their own work. Institutional policies may be either non-existent, unclear, or inconsistent across courses — and may not reflect current student uses of AI (Perkins, 2023; Chan, 2023). As a consequence, many students may not have an established set of ethical guidelines to guide their choices relative to AI-assisted writing — placing them in a position of uncertainty, anxiety, unintentional misconduct or wrongful accusation being wrongfully accused. The potential for wrongdoing via AI detection tools mistakenly classifying legitimate writing produced by non-native English-speaking authors as having been written by an AI tool (Liang et al., 2023) makes these concerns even more critical.

There has been a growing body of research investigating students' views of generative AI in Higher Education. Past research studies have raised concerns related to academic integrity, authorship, dependency, fairness, and proper use (Yusuf et al., 2024; Chan & Hu, 2023). However, most of these studies were based on quantitative and survey data. While surveys are useful for identifying overall attitudes toward AI usage and perceived ethical issues — they offer very little information on how students' reason through ambiguous situations; how students justify their decisions to use AI-assisted writing tools; how students establish an "acceptable" threshold for the degree to which AI assisted them in

completing assignments; and how students respond to ambiguity and uncertainty while producing their writing. Given that most EFL students' uses of AI will be determined not solely by academic pressures but by their need for language support and/or linguistic insecurity -- these limitations are particularly relevant to EFL contexts.

Therefore, the purpose of this study was to explore how EFL university students perceive the moral boundaries surrounding the use of AI tools in academic writing. Additionally, this study sought answers to the following four questions:

1. How do EFL students conceptualize "proper" use of AI in academic writing?
2. What factors shape EFL students' decisions about when to utilize AI-assisted writing tools?
3. What ethical dilemmas or conflicts do EFL students experience when using AI tools?
4. What are the ways in which EFL students resolve conflicts created by AI-assisted writing and compliance with their respective institution's academic integrity policies?

This study is important for a number of practical, institutional policy, and theoretical purposes. Practically, the outcomes from this study will help educators build effective services and courses in writing centers that encourage the responsible application of AI in EFL teaching. Institutionally, these study outcomes will serve as a well-informed basis for institutions to create formalized policies relative to the use of AI-Assisted Academic Writing. From a theoretical perspective, this study contributes to the expanding body of research in Digital Ethics in Language Education through applying Rest's (1986) Four Component Model of Ethical Decision-Making to explore the experiences of EFL Students who apply AI-Assisted Writing. Rest's (1986) model enables consideration of ethical use of AI to be evaluated based upon actions taken and not just whether an individual adheres to their institutions' rules. Further, this study extends previous discussions about EFL education in Saudi Arabia -- a region where the applications of structural models for creating

ethically conscious decision-making processes regarding generative AI have been virtually unexplored (Al-Sofi, 2024; Alshammri, 2024).

## 2. Literature Review

### 2.1 AI-assisted writing in EFL contexts

Large language models appeared in the education field as adaptable tools before they transformed into writing aids. Initial assessments of their educational implications showed substantial prospects, specifically tailored elucidation, immediate feedback, and assistance for learners with varied requirements, alongside persistent issues of output bias, factual inaccuracy, excessive dependence, and the challenges of validating machine-generated content, ultimately determining that the realization of these advantages hinges on intentionally developed skills in both students and educators (Kasneji et al., 2023). Broader capabilities are being utilized in certain ways within instructional contexts of language. The same system is an aid for acquiring content as it is a tool for learning the language of the learner. Questions about the legal implications of using machines arise at multiple levels, including within each individual sentence. ChatGPT has accelerated changes already occurring in second-language (L2) writing with students now beginning to use human-AI collaborative writing. Overviews of language education technology detail a range of potential capabilities that include generating ideas, editing errors, modifying the complexity of the text and providing on demand practice. However, there are warnings from reviewers that for learners to utilize these tools appropriately and efficiently, they will need to develop new digital competencies (Kohnke et al., 2023). As regards

L2 writing, Barrot (2023) identified benefits such as helping to plan; organize and linguistically enhance the writing process, and disadvantages associated with originality, accuracy, and academic honesty. Research often demonstrates significant benefits. Mahapatra (2024), found in a mixed-methods intervention study that the use of ChatGPT significantly improved ESL students' academic writing capabilities. Xiao and Zhi (2023) indicated that EFL students viewed ChatGPT as a useful writing assistant while recognizing the need to critically assess ChatGPT's outputs.

Further investigation on AI-driven evaluation suggests that language models can generate somewhat reliable evaluations of L2 writing and can serve to improve instructor feedback (Mizumoto & Eguchi, 2023).

Warschauer et al. (2023) contend that AI-generated text offers L2 writers an equalizing effect by alleviating the linguistic challenges of composing in English, yet they also identify unresolved dilemmas including the possibility that learners may progress through their education with or without the foundational writing skills their texts appear to exhibit. The distinction between appropriate language assistance and inappropriate content creation is conceptually precarious and vaguely defined for second language writers. This is due to the fact that lexical, syntactic, and organizational support, which were historically the domain of dictionaries, grammar checkers, and writing centers, have now been merged with idea generation within a unified platform. The tension is evident in both Arabic and Saudi EFL settings;

while students appreciate using AIs as assistive technology to support their struggles with vocabulary and grammar as well as organization, they fear dishonesty, over-reliance on an AI tool, and incorrect or false data. In her study titled "Saudi Students' Attitudes towards Using Artificial Intelligence (AI) Tools in Learning English," Al-Sofi (2024) demonstrated how Saudi students see both the advantages of AI use in terms of enhancing their learning speed and improving their ability to learn the English language vs. disadvantages such as relying too heavily on the AI, which may lead to cheating on assignments. Alshammri's (2024) interview-based research has identified two aspects that were positive: students recognize the potential for AIs to provide them with better structure and feedback; however, there are several negative issues that have been identified by students including having fictitious sources referenced in papers, and lack of ability to easily verify information from foreign languages.

The regional results mirror the global trend, however there is a significant caveat for students composing in a foreign language because evaluating AI-generated content constitutes a linguistic challenge in itself, meaning that those who require the most support is the least equipped to analyse it.

## **2.2 Academic integrity and ethics in the AI era**

The capacity of large language models. To produce coherent and apparently unique content has challenged conventional notions of plagiarism, which need a human origin for the

copied material. Perkins (2023) asserts that the utilization of AI by students does not inherently constitute misconduct and that the boundaries of acceptable use must be renegotiated and made explicit in institutional policy. This is because AI-generated content can evade both software detection and the scrutiny of trained individuals. Cotton et al. (2024) characterize the matter as one for preventive institutional measures and advocate for explicit policies that outline the standards for acceptable use and reformed evaluation methods in addition to identification and sanctioning. The inconsistency of detection poses a significant equity issue for L2 populations, Liang et al. (2023) revealed that widely-used GPT detectors inaccurately categorized over fifty percent of authentic essays authored by non-native English speakers as AI-generated, while achieving near flawless accuracy on essays from native speakers. This is due to the detectors' bias against the reduced linguistic variability inherent in L2 writing. Sullivan et al. (2023) found that institutional responses ranged from complete prohibitions to regulated use and transparency requirements. In some media narratives, students are seen mostly as prospective cheats rather than as learners adapting to new technology.

The practical justification for the detection of AI usage in academia has significantly increased. Weber-Wulff et al. (2023) evaluated fourteen detection systems, such as Turnitin on writings authored by humans, generated by AI, paraphrased and machine-translated, concluding that the current tools lack both precision and reliability. Overall accuracy fell below 80 percent.

In addition, performance significantly declined when AI outputs were slightly rephrased or obscured. In conjunction with the previously established bias against non-native writers (Liang et al., 2023), this research indicates that detection is ineffective in various manners and permits machine-generated material to evade detection while identifying genuine L2 compositions. Certain integrity academics advance the notion that the detection paradigm has become outdated. Eaton (2023) suggests that higher education is transitioning into a 'post plagiarism' phase that is characterized by the prevalence of hybrid human–AI writing which necessitates a reestablishment of integrity based on ethics, transparency, and assessment design instead of monitoring textual sources. Regardless of the acceptance of the entire post plagiarism concept, it clarifies the inquiry that this research tackles. If organizations are unable to consistently identify the use of AI, the ethical responsibility of establishing limits significantly transfers to the students.

A frequent observation is that policy has lagged behind practice. Chan (2023) conducted a survey among students, educators, and personnel at Hong Kong institutions. She found considerable uncertainty around allowable practices and suggested an AI Ecological Education Policy Framework that includes educational, governance, and operational aspects. Cross-cultural findings indicate that generative AI may either pose a risk to integrity or serve as a driver for reform (Yusuf et al., 2024). For EFL learners, the uncertainty has increased and a single tool can serve as valid language assistance in one class

while being deemed forbidden content creation in another, requiring students to discern the limits independently.

### **2.3 Students' viewpoints and choices in decision-making**

Uncertainty frequently emerges in research regarding students' perceptions and viewpoints. Chan and Hu (2023) conducted a survey including 399 university students and discovered predominantly favorable perceptions of using generative AI. However, the findings do show worries regarding related issues such as accuracy, ethics and impacts on students' personal growth. Qualitative research enhances this understanding, in an L2 writing practicum Yan (2023) discovered that learners appreciated ChatGPT's efficiency, however, they also expressed unprompted worries over academic integrity and educational fairness particularly the disparity between users and non-users. Research involving interviews with Saudi EFL students has indicated a duality in their appreciation for structural and feedback assistance, coupled with apprehensions regarding dishonesty, excessive dependence, and falsified references (Alshammri, 2024). L2 writers have voiced apprehensions regarding the decline in learning and the diminishing of authorial voice, urging institutions to establish definitions for effective and ethical usage of AI (Yan, 2023). Some research adopts a moral psychology perspective instead of an attitude assessment approach to student behavior. Bandura (2016) posits that moral disengagement refers to the cognitive processes (euphemistic labelling, advantageous comparison, diffusion of responsibility, and displacement of responsibility

onto circumstances or authorities) that enable individuals to act in manners they would typically denounce while preserving their moral self-perception. This perspective predicts which patterns will be visible in writing supported by AI ("help requested," "expression improved") describing the creative process; normalizes use via references to commonality ("everybody does it"); and delegates the issue of boundaries (i.e., acceptable uses) to schools/educators/institutions. In theory, these methods should fit into Rest's components illustrating how, without experience with wrongdoing, both sensitivity and motivation may decrease; also providing an outline of the normalization and justification processes that research has only suggested existed.

Collectively, this study suggests that student use of AI is influenced by a variety of factors; some practical (e.g., time constraints, the level of difficulty of tasks assigned, self-assurance regarding English), others social (e.g., perceived expectations from teachers, norms among peers, individual views on authorship/originality). Nonetheless, most existing studies evaluate opinions through surveys instead of investigating how individual students handle uncertain scenarios and justify their choices or how they may feel guilt and uncertainty throughout the writing process. The psychological aspects of AI use, including normalization and rationalization, are still insufficiently examined, especially among EFL populations where the difference between language support and content generation is

inherently unclear (Warschauer et al., 2023).

## 2.4 The conceptual framework

The study was based on Rest's (1986) Four-Component Model. This model suggests that there are four types of cognitive processing that lead to morally correct behavior. These include: Moral sensitivity refers to recognizing a situation involving conflict over what is right; moral judgment refers to evaluating the situation to determine what should be done; moral motivation involves prioritizing moral values over competing personal or practical interests; and moral character/action refers to having the persistence and ability to carry out the chosen ethical action. Applying this model for AI-assisted writing offers a perspective to enquire, do students regard the usage of AI as an ethical concern? How do they evaluate specific applications as permissible or impermissible? How do they balance integrity with efficiency and linguistic uncertainty? How do they respond to those evaluations within time constraints and ambiguous policy? The framework aligns with the interpretivist perspective of this research as each component is developed through the student's personal interpretation rather than a definition of wrongdoing imposed from outside.

Rest's framework was selected above other credible alternatives for two primary reasons. Digital-literacies frameworks are beneficial for outlining the skills required for AI utilization (Kohnke et al., 2023), but regarding ethical behaviour, ethics is only one of several competencies and offer insufficient insights into the psychological mechanisms that influence a student's transition from identifying an ethical

dilemma to taking action. The research enquiry of this study specifically focusses on that process. The main distinction is as follows. Stage theory of moral development views individuals in terms of a developmental continuum; The Four Component Model does not view its students as hierarchical. Rather than providing an example where students are sensitive but have underdeveloped judgment or demonstrate strong judgment but poor implementation, it provides an opportunity for researchers to examine the stage at which the dilemma occurs. Additionally, the four-component model serves as a guide for the analytical process used within this study. For instance, each of the inquiries (RQ1 -- Sensitivity & Judgement; RQ2 -- Motivation; RQ3 -- Conflict amongst Components; RQ4 -- Character when faced with Institutional Pressure) can be related back to the above research question.

## 2.5. Research Gap

The present body of literature indicates the advantages that AI writing tools may have for EFL students (Mahapatra, 2024; Barrot, 2023); the disruption to traditional models of academic integrity (Perkins, 2023; Cotton et al., 2024); and the ambivalent views that many learners have toward using them (Yan, 2023; Chan & Hu, 2023). There are three challenges to this area. Firstly, the base of evidence has primarily consisted of large-scale, quantitative surveys and educator/policy-maker viewpoints with limited qualitative data on how students' transverse in the use of technology. Second, research on the ethics of AI-assisted writing generally documents the existence of ethical concerns but provides limited insight into how students interpret these

concerns, negotiate ambiguous situations, or determine appropriate responses (Al-Sofi, 2024; Alshammri, 2024). Relatively little research has applied a clear ethical decision-making framework to students' real-life experiences with AI-assisted writing.

## 3. Methodology

### 3.1 Study Design

The current study utilized a qualitative, exploratory research design based upon an interpretivist social constructivist perspective. Students' ethical reasoning related to AI assisted academic writing is both subjectively perceived and dependent upon contextual influences including interaction with technology, peers, instructors and expectations generated by institutions. A qualitative research methodology was deemed most suitable to capture the complexities and meanings that are individually constructed from such experiences. The current study aimed to explore participants' perception, reasoning and experience relative to their views of AI usage in academic writing tasks as opposed to developing a theoretical model or providing an extensive case-based analysis of a particular bounded setting.

### 3.2 Participant Recruitment and Sampling

Through purposeful sampling six Saudi English as a Foreign Language Master's degree students from the same university participated in this study. In order for eligibility criteria to be met participants had to be registered in a Master's program in EFL, have used at least one AI supported academic writing tool for an academic writing task during the preceding academic term and agree to participate in discussions about their experiences. Participants were assigned participant numbers for purposes of maintaining anonymity and all data were collected over five interviewing sessions between July 27th and August 3rd, 2026, one session involved two participants, while the remaining four sessions were conducted individually.

### 3.3 Instrumentation

Data collection utilized a semi-structured interview guide developed consistent with the objectives of the study, the research questions, and Rest's (1986) Four Component Model of Ethical Decision Making. The instrument included five major sections: (a) participants' background information and knowledge about AI supporting tools; (b) perceptions of ethically acceptable uses of AI supporting tools in academic writing, which addressed Question 1; (c) Factors that influenced participants' decisions to employ AI supporting tools for academic writing, which addressed Question 2; (d) Ethical concerns and dilemmas that arise when employing AI supporting tools for academic writing, which addressed Question 3; and (e) Perceptions of the level of expectation among universities regarding institutional academic integrity in relation to AI supported academic writing, which addressed Question 4. Each session concluded with additional open-ended questions to allow participants to reflect upon aspects of their experiences they found important.

Each section of the interview guide served as a flexible outline rather than a rigid scripted format. Follow-up questions and probes were used to obtain specific examples from participants and to obtain detailed descriptions of participants' experiences.

### 3.4 Procedures

Interviews lasted approximately 30–45 minutes per participant. Prior to each interview, participants received written details regarding the study's objective, the voluntary aspect of their involvement, the intended use of the data, and their rights to withdraw from participation. Written documentation confirming participants' informed consent was obtained prior to commencing audio recordings.

Interviews were conducted primarily in English, although participants were permitted to use Arabic when necessary to express their views accurately. The interviewer maintained a neutral stance toward AI use and explicitly informed participants that their responses would remain

confidential, would not be disclosed to their instructors, and would not be used as evidence of academic misconduct. This assurance was intended to encourage open discussion of potentially sensitive experiences. Following the interviews, the audio recordings were transcribed and the transcripts were checked against the original recordings for accuracy before analysis.

### 3.5 Data Analysis

The interview transcripts were analysed thematically following Braun and Clarke's (2006) six-phase approach: familiarisation with the data, initial coding, generation of candidate themes, review of themes, definition and naming of themes, and production of the final report. A hybrid deductive–inductive approach was used. Deductive coding provided an initial organisational framework based on the four research questions and relevant components of Rest's (1986) Four-Component Model. Inductive coding techniques were also used as part of the larger thematic categories to allow for pattern identification and theme development from the participants' experiences. The use of inductive coding approaches allowed researchers to maintain theoretical understanding of the research questions while maintaining sensitivity to participant experiences and perspectives that may have been unaccounted for within a predetermined conceptual model. Quotations from participants are contained in Section 4; all editing (i.e., grammar and readability) was limited to the extent necessary to maintain the participant's original intent.

### 3.6 Trustworthiness and Ethical Considerations

The trustworthiness of this study was assessed using Lincoln and Guba's (1985) four main criteria for trustworthiness: credibility, transferability, dependability and confirmability. Credibility was achieved by member checking the interpretations of the participants' data and also via peer debriefing. Transferability was furthered through providing a clear and detailed description of both the study context and participants, along with all data collection processes. Maintaining an

audit trail of all methodological and analytic decisions provided evidence of dependability. Finally, through keeping a reflexive journal that monitored the researcher's own assumptions, specifically regarding her role as an EFL teacher, the potential for bias was minimized and therefore confirmability was enhanced.

The ethical considerations were also taken into account at every stage of this study. Participation was voluntary; participants gave informed consent prior to collecting any data; participants were made aware of how the data collected during the study would be utilized; all identifying information was removed from all audio recordings and written transcripts; participants were allowed to withdraw from the study at any time after they participated in the study. All of the participant's audio recordings and written transcripts were safely stored so that only members of the research team had access to them.

#### 4. Results

The results were structured based on the four research questions, with themes created inductively from the interview data. A contextual theme related to participants' experiences with academic writing in English is shown first, as it frequently appeared throughout the interviews. Participant quotations are used to illustrate the identified themes.

##### 4.1 Context of EFL Academic Writing

Participants described difficulties arising from the difference between the informal English encountered in everyday life and the formal language required for academic writing. P1 stated, "It might be hard if it is academic, because we are used to the informal language from social media and from the series and movies we watch. The language barrier has a really significant impact on our academic writing." She identified writing and speaking as particularly challenging skills and also referred to difficulties with academic genre structure: "If I don't fully understand the structure of a research paper, the literature review, the methodology, the results and the

discussion, I may have good ideas, but I might not present them in the correct academic way."

Participants differed in their perceptions of writing difficulty. P2 reported, "For me, writing is much easier than speaking, because I can organise my thoughts on paper." P1 and P2 also described moving between Arabic and English during composition. P1, for example, reported switching languages "to see if I have the suitable word" when she could not immediately express an idea in English.

Lack of formal writing support was also reported. P5 explained, "When I was an undergraduate student, I never got an academic writing centre to understand my writing, whether it's good or not. So now, when we use AI, it's like a replacement for a writing centre." Participants also described strategies that did not depend entirely on AI. P1 explained how she prepared vocabulary before timed writing tasks: "Before an exam I prepare myself by searching for suitable academic verbs, like demonstrate, and even the conjunctions, so I can use them in my writing." P4 indicated that he limited his AI use to "summarising and searching information only."

##### 4.2 RQ1: Students' Understanding of Ethical AI Use

Participants did not always recognize the term 'ethical AI use' as a known term. P1 said, "Ethical use...You're saying the right way to use AI?" P4 claimed he had never heard of "ethics" associated with AI, however, after being provided several examples, participants were normally able to tell what kinds of uses they deemed acceptable, versus those they did not.

Three different ways to define "Ethical AI Use" arose.

First, the difference between aid and substitution was identified by both P1 and P2. P1 said, "I could use AI to assist me in developing ideas for a paper, or to help me communicate those ideas; I would not trust AI enough to complete the writing of the paper." P2 stated, "Students should utilize AI to enhance their learning; students should not allow themselves to substitute AI for

critical thought, nor should students allow themselves to 'cheat' using AI."

Second, concerns regarding the extent of dependence upon AI were expressed. For example; P5 defined an appropriate amount of use for himself (numerically): "I do not depend completely upon the AI. For example, I use it about 20% — so I'm going to check grammar and syntax to make sure what I've written makes sense; or occasionally, I'll ask the AI to review my essays to determine if I've overlooked anything within my research and analysis. I'd classify this as unacceptable use if I were to base an entire paper solely on the AI's response." P2 stated she thought the instructor guidelines provided by her institution were relatively clear. However, P3 indicated that clarity was related directly to the application's privacy policies. P1 described the boundary of where acceptable and unacceptable AI use exist as being contingent upon the type of academic tasks.

#### **4.3 RQ2: Factors Influencing Decisions About AI Use**

Participants chose whether or not to use AI-based writing tools due to numerous factors.

For many participants, one factor was linguistic insecurity, including how they utilized their AI to aid them in utilizing vocabulary in an essay, while still being able to express themselves in their own voice.

P1 said, "I don't want to write using vocabulary I have never seen before. If I want this to be my essay, and my words, and I want to choose vocabulary that fits me, I will just choose vocabulary that is familiar to me." Similarly, P2 said that she uses AI to help her translate her thoughts from Arabic to English (and) "to create sentences and pick vocabulary to make it clear what I am thinking," and added that she writes differently depending on who the audience is.

Additionally, some participants reported that time constraints influenced their choice of whether or not to use AI-based writing tools. Deadlines from assignments combined with limited time available to complete those

assignments all contributed to the likelihood of utilizing AI. In addition to time pressures, difficulty with task completion (i.e., perceived difficulty with writing), complexity of topic, and accessibility of information were all factors mentioned by participants. For example, P5 reported that he would be more likely to use AI when working on difficult topics or when he could not find enough resources. Additionally, participants' perceptions of AI accuracy and efficiency of AI was a factor as well. For example, P3 said: "The work is certainly efficient and motivating; in fact, it is quite accurate as well. Sometimes when I am working independently, I make mistakes. However, AI is generally very close to being accurate."

Instructor expectations also played a role in decisions made about AI use; however, the expectations of instructors varied significantly. As such, P5 stated: "There are teachers who allow their students to utilize AI and there are teachers who tell their students not to utilize AI. These instructors want to assess whether or not they can see the student's genuine work rather than the AI generated work." On the other hand, P6 stated that the expectations of their instructors did not impact how often he/she utilized AI.

#### **4.4 RQ3: Ethical Dilemmas and Tensions**

Participants' accounts identified three major concerns: uncertainty over what constitutes acceptable AI use, detection issues related to AI-generated content, and reliability and trustworthiness in the generation of AI content.

Concern one was a lack of clarity around boundaries. Participant P1 noted, "I love the support of AI, but I have no idea how much support I can get. I have no idea where my limits lie."

Concern two was the use of AI-detection and similarity tools. P1 used these types of tools and modified her writing based on output from them.

"I found out when you type your responses extremely well, those detector programs will believe you're an AI. And when I respond poorly, they believe me to be human. When I convey my thoughts strongly, then those programs indicate

I'm using an AI. Thusly, I changed my approach to writing; I was quite astonished by this."

P2 recalled an instance from her past in which plagiarism detection software flagged similarities in both her own independent composition and her own content. She stated she relied upon her own thinking while creating the paragraph, but there were ideas that appeared plagiarized and coloured red as if she had copied them. As a result of this experience, she has been using plagiarism-detection software to verify her work prior to submitting. The third issue was how reliable is AI generated material? (especially referencing and factual data). P5 said, "The AI could never provide me with the proper references. And some of those references aren't even legitimate. It may possibly create new ones. Therefore, I need to be able to confirm each reference and each study." Likewise, P4 expressed his concerns over whether AI generated content and translations were indeed representative of what he asked for. They do so constantly due to the uncertainty of being within an ambiguous gray area."

#### **4.5 RQ4: Reconciling AI Use with Academic Integrity Expectations**

No participant identified a written institutional policy regulating the use of generative AI in their academic writing.

P5 said, "I didn't see anything about using AI, if it's allowed or forbidden. I couldn't find anything like that. It has to be verbal. I've never seen anything like this." Generally speaking, participants described instructors as the primary authority guiding them. P2 said, "Professors teach us the ethics of writing before teaching the act of writing and what can be accepted in the course and what should not be done by students."

Although P4 had some knowledge of departmental policies regarding unacceptable use, he described these policies as minimal. When questioned on ways participants maintained their own perceptions of unacceptable uses of AI, participants discussed multiple methods. Participants used methods such as comparing the work of students against plagiarism or AI

detection software; verifying AI sourced information, changing AI produced content into content they were familiar with; and limiting their use of AI to later phases of the writing process. Participants expressed a preference for clearer institutional guidance. P5 stated that universities should "make a clarification of how we're going to use AI" and provide more systematic expectations so that students understand the permissible level and purposes of AI assistance.

Overall, the findings showed substantial variation in how participants defined unacceptable AI use, why they used AI, and how they attempted to comply with academic integrity expectations. Across the interviews, the most recurrent findings were uncertainty about unacceptable boundaries, the use of AI for linguistic and pragmatic support, inconsistent instructor guidance, concerns about detection and accuracy, and a preference for clearer institutional rules.

## **5. Discussion**

### **5.1 Students' Improvised Ethical Frameworks for AI Use**

The findings indicate that participants were attempting to construct their own ethical frameworks for AI-assisted academic writing in the absence of a clearly identified institutional standard. Their judgments drew on previous instructor comments, application privacy requirements, self-imposed limits on AI use, and feedback from AI-detection or similarity tools. These findings extend Rest's (1986) Four-Component Model in several ways.

First, even though many students acknowledged there were potential ethical issues with using AI for assisting with writing, those ethical issues were often inconsistent and did not result from clear institutional definitions of what constitutes appropriate uses of AI.

Secondly, while the participants in this study had differing views about how they would define "ethically" using AI to help them write, one common thread in all of their responses was that AI could be used to support academic writing so long as it does not take away the student's own intellectual contribution.

ensuring that students understand the content they submit rather than relying substantially on AI-generated text

## 5.2 Why EFL Writers May Experience the Boundary Differently

The results of the study also demonstrated the importance of the EFL context when analyzing students' perception of the extent to which they find it acceptable to receive assistance from AIs. Warschauer et al. (2023) have reported that generative AI may reduce the linguistic barriers for second language writers, while creating greater ambiguity in determining where language support begins and ends and where substantive authorship is created. These results confirm that tension. Many of the participants viewed AI primarily as enhancing vocabulary, phraseology, register and developing ideas that were generated by the students themselves. P1's preference for using vocabulary "those suits" her, and P2's view that AI helps "communicate Arabic thought in English," suggest that there is a close relationship between AI assistance and self-expression through language rather than just generating content. As such, the ethical decision-making process concerning AI in the classroom was not necessarily whether AI should be utilized at all, but how much language mediated through AI could be used before the final product would no longer be considered authored by the student.

P5's account provides another layer of complexity to the institutional aspect of this debate. He views AI as partially replacing the academic writing support he had never previously received. These comments add nuances to previous research that generative AI supports EFL learner in vocabulary development, organizational skills, and overall writing performance (Xiao & Zhi, 2023; Mahapatra, 2024). For many learners, AI may serve as an accessible means of receiving language support and therefore regulations restricting AI usage need to consider the educational goals that students are seeking to achieve via these technologies. The results of this study also reflect similar reports from researchers investigating AI assisted learning within Saudi educational environments regarding students' positive attitudes toward AI assisted learning and their concern regarding its perceived ethics/educational value (Al-Sofi, 2024;

Alshammari,

2024). The current qualitative results enhance this literature by illustrating how this uncertainty is felt and navigated in the process of academic writing.

## 5.3 Policy Ambiguity and Relying On AI

There are many studies that demonstrate ambiguous or inconsistent institutional policies create uncertainty regarding the use of generative AI in Higher Education (Perkins, 2023; Chan, 2023), in addition to creating uncertainty for students' confidence in their own decision-making processes, the results suggest ambiguity may lead to varying forms of student-regulated AI usage strategies. Students indicated development of personal rules when formal guidance was not available. Examples of these were self-imposed numerical limits on the extent of AI use; differentiating between type of assignments (e.g., essays vs. multiple-choice questions); using limitations related to specific AI functions (e.g. only allowing AI to assist with grammar corrections and never paraphrasing); and trusting the expectations provided by instructors. P5's observation that some students rely heavily on AI simply because they are operating in a "grey zone" indicates that ambiguous expectations may result in inconsistent or over-reliance upon AI. This should be interpreted with caution since it is solely based on the perceptions of participants and does not provide direct evidence of causality. An additional dimension of understanding this occurrence is habitual usage. A large number of participants stated that due to their heavy reliance on AI in both daily academic and personal tasks, utilizing AI was no longer a conscious decision; it had simply become routine. Therefore, when considering the study of the use of AI as merely an independent choice or as one singular act, the focus should be shifted from whether or not students utilize AI, to what reasons prompted them to use AI, how they incorporate the output generated by AI in their work, and whether or not disclosure of such occurs.

## 5.4 AI Detection and the Experience of EFL

**Students** The primary finding was that users responded to both an automated detection system and a similarity system. One (P1) of the users stated she made changes to her writing after realizing strong expressions in English were sometimes automatically classified as having been created using artificial intelligence. The single

case is illustrative of a potential negative behavioural response when students rely on such tools. The concern expressed above is supported by a larger body of literature calling into question the ability of AI-detection systems to accurately identify content generated by humans (Weber-Wulff et al., 2023). Additionally, Liang et al. (2023) found significantly high rates of false positives when detecting writing created by individuals who write in languages other than English. The current study provides a qualitative viewpoint through describing how participants perceive unreliable detection systems will impact their writing behaviours. In P1's case, the reaction was not simply fear, but the intentional reduction of genuine writing. These types of experiences have significant implications for policy concerning academic integrity. As students alter their genuine writing styles primarily to avoid detection by algorithms, detection systems may inadvertently cause institutions to move the emphasis of academic integrity issues away from authorship processes, and toward producing content that is sufficiently "human" to software. This supports arguments that academic integrity should not primarily depend on the use of automated AI detection (Eaton, 2023). P2's previous experience with similar-detection software also indicates that concerns about the use of automated classification exist well beyond generative AI. However, the fact that P2 now checks her original work prior to submitting it demonstrates how having experienced being flagged by a detection system caused her to engage in defensive self-monitoring.

### 5.5 Limits of Individual Self-Regulation

Findings also show the variability in how participants perceived their own AI usage. Some participants indicated varying levels of uncertainty/ anxiety regarding their use of AI. Others had virtually no indication of ethical concern regarding their usage because they believed their usage was correct. For instance, participant P5 stated he used AI appropriately in his writing, however, he acknowledged that overuse could impede the writing abilities of his peers. Accounts similar to those above can be explained within the context of Bandura's (2016) ideas about individual self-regulatory behaviours related to morality and the processes by which an individual separates their actions from behaviours that are viewed as less desirable. The explanations provided above should be interpreted cautiously since the interviews

conducted did not assess specific mechanism(s) associated with moral disengagement. Participants P3 and P4 illustrated additional examples of how students may view the ethics of using AI in a completely different conceptual manner when assessing their own understanding of the issue. These varied ways of thinking illustrate another challenge that exists with depending solely on individual self-regulation. As noted previously, if there is a lack of common terms or expectations articulated by institutions, students may make internal consistency decisions based on differing definitions of "ethical" use of AI.

**5.6 Implications for Theory** There are two main implications that result from the study results related to the implementation of Rest's (1986) Four-Component Model to assistive technology-assisted writing. One major implication relates to whether or not habituated AI use affects the component of moral sensitivity in Rest's (1986) Four-Component Model. Typically, one would assume that an individual recognizes that a given situation demands an ethical response prior to developing any level of moral sensitivity. In many cases, the habituation of AI into daily writing routines has led to situations where use occurs without conscious realization that it involves making an ethical decision. A second major implication is that moral judgment seems influenced at least partially by the existence of a clearly defined external criterion. Although participants could typically indicate extremes, i.e., submitting assignments generated by AI alone, they found difficulty in determining moderate degrees of use. Thus, institutional ambivalence may limit the extent to which an individual makes ethical judgments. As noted earlier, this study suggests that each component of the decision-making process can be examined in conjunction with an individual's learning environment. Academic environments provide opportunities for educators to influence the degree to which students are able to recognize, assess, and respond to ethical issues.

**5.7 Implications for Policy and Educational Practices** Study results support calls for clearer and more consistent institutional guidance regarding use of generative AI technologies (Perkins, 2023; Chan, 2023). Policies need to transition away from broad declarations that simply permit or prohibit AI and provide more detailed descriptions of permissible versus

impermissible use according to the functions served by the tools.

Examples include establishing distinctions between brainstorming, language editing, rewriting sentences/paragraphs, translating words/phrases, extracting summaries, finding sources, verifying facts, providing structural help and generating content. The distinction established by the functional nature of the tool assists participants who attempt to differentiate types of uses similar to those identified in this study. A functional perspective might prove more effective than imposing arbitrary percent limitations.

In addition to providing consistent policies across all courses in an institution, guidelines should allow for instructors to establish assignment-specific requirements. Many participants relied on oral directions from individual faculty members resulting in potential contradictions in expectations. Providing written and understandable guidelines could help minimize inconsistencies due to variations in course expectations. The findings further suggest a need for explicit instruction about the limitations of generative AI, including fabricated references, inaccurate information, and the inability of AI systems to independently guarantee factual reliability. Students should also be given clear procedures for disclosing legitimate AI assistance where disclosure is required. None of the participants described routinely documenting prompts or formally reporting AI use. AI-detection tools should be used cautiously and should not constitute the sole basis for allegations of misconduct. Where questions about authorship arise, greater emphasis may be placed on transparent writing processes, drafts, oral explanation, formative assessment, and assignment design.

Finally, institutions should consider why students turn to AI. Where AI is being used to compensate for limited academic writing support, restrictive policies alone may remove an important source of assistance without addressing the underlying educational need. Writing centers, EFL support services, formative feedback, and explicit instruction in academic writing may therefore

form an important part of responsible institutional AI policy.

### 5.8 Limitations and Future Research

Several factors need to be taken into account when evaluating the study's findings. First, this study is limited by its small sample size of only six participants drawn from a single Saudi post-secondary educational institution. While this study employed an interpretive design, which provided rich contextualized narratives (i.e., stories) about each student's experience, it did not provide a statistical representation of the opinions of all EFL students studying in Saudi Arabia. Thus, caution needs to be exercised when generalizing the results of this study to other EFL students studying in Saudi Arabia or elsewhere. Second, while participants for this study were identified through their previous use of AI-supported writing tools, their perspectives may differ significantly from those students who do not utilize such technology, have no or limited access to AI writing tools, or have gone through formal institutional disciplinary actions involving AI. Third, this study utilized self-reported interview data for a potentially sensitive topic. Participants' reports may have been influenced by either recollection difficulties or social desirability bias despite assurances of confidentiality. Fourth, as per the request of two participants, interviews were conducted jointly with two participants. It is possible that another participant's presence may have impacted the level or type of disclosure made by participants. Finally, whereas this study focused upon the perspectives of the participants, it did not independently verify whether they used AI in their submissions, how much instructor support they received, or the effectiveness of the AI-detection tools that they referenced. Therefore, results regarding detection should be viewed as students' reported experiences and not necessarily as independent evaluations of specific detection systems.

### 6. Conclusion

This research investigated how Saudi EFL graduate students understood the ethical borders of using AI-generated assistance while developing their own academic papers. While

there were many different standards of what constituted acceptable use of AI, most students relied upon some combination of their instructors' approval, personal ethics/philosophy, concern for privacy, establishing their own boundaries of responsibility, detecting whether an assignment contained similar wording, etc. Regardless of the different methods used to evaluate AI, one consistent finding among the students was a distinction between using AI as a "tool" to assist them in completing assignments (and thus allowing them to maintain control over their work) versus permitting AI to replace their own original ideas/concepts in producing written products. The use of AI was also affected by factors such as the uncertainty of language proficiency, time constraints, task difficulty,

perceived usefulness/effectiveness of using AI, requirements from teachers and traditional practices in class. These findings indicate a need for universities and colleges to develop clearly articulated and uniformly applied policy documents that detail allowable uses of generative artificial intelligence tools. In addition to providing these policy documents, faculty are encouraged to provide students with instruction on the limitations and potential flaws of content produced through generative AI. Additionally, institutions should provide adequate support to EFL writers (in terms of assistance with the development of their writing skills). Finally, automated plagiarism detection software should not be the sole basis for determining academic misconduct.

## 7. References

- Al-Sofi, B. B. M. A. (2024). Artificial intelligence-powered tools and academic writing: To use or not to use ChatGPT. *Saudi Journal of Language Studies*, 4(3), 145–161. <https://doi.org/10.1108/SJLS-06-2024-0029>
- Alshammri, A. (2024). ChatGPT in English writing: Experiences and perceptions of Saudi EFL students. *Amazonia Investiga*, 13(78), 47–57. <https://doi.org/10.34069/AI/2024.78.06.4>
- Bandura, A. (2016). *Moral disengagement: How people do harm and live with themselves*. Worth Publishers.
- Barrot, J. S. (2023). Using ChatGPT for second language writing: Pitfalls and potentials. *Assessing Writing*, 57, Article 100745. <https://doi.org/10.1016/j.asw.2023.100745>
- Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77–101. <https://doi.org/10.1191/1478088706qp063oa>
- Chan, C. K. Y. (2023). A comprehensive AI policy education framework for university teaching and learning. *International Journal of Educational Technology in Higher Education*, 20, Article 38. <https://doi.org/10.1186/s41239-023-00408-3>
- Chan, C. K. Y., & Hu, W. (2023). Students' voices on generative AI: Perceptions, benefits, and challenges in higher education. *International Journal of Educational Technology in Higher Education*, 20, Article 43. <https://doi.org/10.1186/s41239-023-00411-8>
- Cotton, D. R. E., Cotton, P. A., & Shipway, J. R. (2024). Chatting and cheating: Ensuring academic integrity in the era of ChatGPT. *Innovations in Education and Teaching International*, 61(2), 228–239. <https://doi.org/10.1080/14703297.2023.2190148>
- Eaton, S. E. (2023). Postplagiarism: Transdisciplinary ethics and integrity in the age of artificial intelligence and neurotechnology. *International Journal for Educational Integrity*, 19(1), Article 23. <https://doi.org/10.1007/s40979-023-00144-1>
- Kasneci, E., Seßler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., Gasser, U., Groh, G., Günnemann, S., Hüllermeier, E., Krusche, S., Kutyniok, G., Michaeli, T., Nerdel, C., Pfeffer, J., Poquet, O., Sailer, M., Schmidt, A., Seidel, T., ... Kasneci, G. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. *Learning and Individual Differences*, 103, Article 102274. <https://doi.org/10.1016/j.lindif.2023.102274>

- 
- Kohnke, L., Moorhouse, B. L., & Zou, D. (2023). ChatGPT for language teaching and learning. *RELC Journal*, 54(2), 537–550. <https://doi.org/10.1177/00336882231162868>
- Liang, W., Yuksekgonul, M., Mao, Y., Wu, E., & Zou, J. (2023). GPT detectors are biased against non-native English writers. *Patterns*, 4(7), Article 100779. <https://doi.org/10.1016/j.patter.2023.100779>
- Lincoln, Y. S., & Guba, E. G. (1985). *Naturalistic inquiry*. Sage.
- Mahapatra, S. (2024). Impact of ChatGPT on ESL students' academic writing skills: A mixed methods intervention study. *Smart Learning Environments*, 11, Article 9. <https://doi.org/10.1186/s40561-024-00295-9>
- Mizumoto, A., & Eguchi, M. (2023). Exploring the potential of using an AI language model for automated essay scoring. *Research Methods in Applied Linguistics*, 2(2), Article 100050. <https://doi.org/10.1016/j.rmal.2023.100050>
- Perkins, M. (2023). Academic integrity considerations of AI large language models in the post-pandemic era: ChatGPT and beyond. *Journal of University Teaching and Learning Practice*, 20(2), Article 07. <https://doi.org/10.53761/1.20.02.07>
- Rest, J. R. (1986). *Moral development: Advances in research and theory*. Praeger.
- Sullivan, M., Kelly, A., & McLaughlan, P. (2023). ChatGPT in higher education: Considerations for academic integrity and student learning. *Journal of Applied Learning and Teaching*, 6(1), 31–40. <https://doi.org/10.37074/jalt.2023.6.1.17>
- Warschauer, M., Tseng, W., Yim, S., Webster, T., Jacob, S., Du, Q., & Tate, T. (2023). The affordances and contradictions of AI-generated text for writers of English as a second or foreign language. *Journal of Second Language Writing*, 62, Article 101071. <https://doi.org/10.1016/j.jslw.2023.101071>
- Weber-Wulff, D., Anohina-Naumeca, A., Bjelobaba, S., Foltýnek, T., Guerrero-Dib, J., Popoola, O., Šigut, P., & Waddington, L. (2023). Testing of detection tools for AI-generated text. *International Journal for Educational Integrity*, 19(1), Article 26. <https://doi.org/10.1007/s40979-023-00146-z>
- Xiao, Y., & Zhi, Y. (2023). An exploratory study of EFL learners' use of ChatGPT for language learning tasks: Experience and perceptions. *Languages*, 8(3), Article 212. <https://doi.org/10.3390/languages8030212>
- Yan, D. (2023). Impact of ChatGPT on learners in a L2 writing practicum: An exploratory investigation. *Education and Information Technologies*, 28(11), 13943–13967. <https://doi.org/10.1007/s10639-023-11742-4>
- Yusuf, A., Pervin, N., & Román-González, M. (2024). Generative AI and the future of higher education: A threat to academic integrity or reformation? Evidence from multicultural perspectives. *International Journal of Educational Technology in Higher Education*, 21, Article 21. <https://doi.org/10.1186/s41239-024-00453-6>
-