AI Detection Analysis Report

Generated: June 27, 2026 | Topic: Use of artificial-intelligence tools and learning motivation among undergraduate students: a quantitative survey study

Client Action Summary

Overall AI-detection risk: Not determined — the AI-detection model did not run on this paper (it is unavailable or its access token is not configured); only the heuristic indicators apply. See the note below the summary table for details.

Confidence: Low — based on local heuristic indicators only.

Sections to review: Research Questions

Recommended actions:

- Vary sentence length and openings; make the phrasing more specific and source-based.
- Replace with a specific, content-based opening.
- Reduce repeated transitions — vary or remove some occurrences.
- Vary how sentences begin.

Section-by-Section Action Plan

Section	Status	What to improve
Abstract	**OK (normal for this section type) — Low sentence variation is normal here.**	No change needed.
Introduction	**OK**	No change needed.
Core Concepts and Definitions	**OK**	No change needed.
Review of Prior Research	**OK**	No change needed.
Synthesis and Research Gap	**OK**	No change needed.

Section	Status	What to improve
Research Questions	Review	Vary sentence length and openings; make the phrasing more specific and source-based.
Research Hypotheses	OK (normal for this section type) — Low sentence variation is normal here.	No change needed.
Research Design	OK	No change needed.
Population and Sample	OK	No change needed.
Research Instruments	OK	No change needed.
Procedure	OK	No change needed.
Data Analysis	OK	No change needed.
Ethical Considerations	OK	No change needed.
Characteristics	OK	No change needed.
Descriptive Statistics	OK	No change needed.
Hypothesis Testing	OK (normal for this section type) — Low sentence variation is normal here.	No change needed.
Interpretation of key findings	OK	No change needed.
Comparison with the literature	OK	No change needed.
Contribution and practical implications	OK	No change needed.
Limitations	OK	No change needed.
Future Research	OK	No change needed.
Summary of Main Findings	OK	No change needed.
Conclusions	OK	No change needed.
Recommendations	OK	No change needed.

Section	Status	What to improve
References	**OK (normal for this section type) — Low sentence variation is normal here.**	No change needed.

Phrasing to Revise

Issue	Suggested fix
Formulaic opening: “in this study” (appears 2 times)	Replace with a specific, content-based opening.
Overused transition: “overall” (appears 4 times)	Reduce repeated transitions — vary or remove some occurrences.
Many sentences begin with “the” (67 times)	Vary how sentences begin.
Many sentences begin with “this” (28 times)	Vary how sentences begin.
Many sentences begin with “for” (8 times)	Vary how sentences begin.
Many sentences begin with “as” (5 times)	Vary how sentences begin.

Executive Summary

This report analyses the document for AI-generated content indicators, comparing scores **before** and **after** the academic humanisation (polish) stage.

Metric	Before Polish	After Polish	Change
AI Detection Score	N/A	N/A	N/A
Burstiness	79%	78%	**↓ 1pp**
Perplexity Proxy	78%	78%	**↓ 0pp**

⚠ AI Detection Score is unavailable: the HuggingFace API token (`HF_TOKEN` / `HUGGINGFACE_TOKEN`) is not configured on this deployment. Only the local heuristic metrics (Burstiness, Perplexity Proxy) are shown above. Set the token to enable the chatgpt-detector-roberta classifier — it is the most reliable single indicator in this report.

Understanding the Metrics

AI detection tools are not infallible. Different detectors use different methods and may produce contradictory results. This report provides indicators based on statistical analysis and a dedicated detection model, but results should be interpreted as guidance, not absolute truth.

AI Detection Score (chatgpt-detector-roberta model)

What it measures: This score is produced by a dedicated AI detection model (Hello-SimpleAI/chatgpt-detector-roberta), a RoBERTa transformer fine-tuned specifically to classify text as human-written or AI-generated. The model analyses linguistic patterns, word choice distributions, and contextual coherence to produce a probability score.

How to read it: 0% = certainly human-written, 100% = certainly AI-generated — the higher the score, the more likely the text is to be flagged as AI-generated. The risk levels below explain what each band means.

Risk Levels

Low Risk (0–30%): The text exhibits characteristics typical of human writing. Most AI detectors are unlikely to flag it.

Medium Risk (30–60%): The text has some AI-like patterns. Some detectors may flag portions. Consider reviewing flagged sections for natural phrasing.

High Risk (60–100%): The text shows strong AI-generation signals. Most detectors will likely flag it. Significant rewriting may be needed.

Burstiness

What it measures: Burstiness measures the variation in sentence lengths throughout the document. Human writers naturally produce “bursty” text — a short punchy sentence followed by a long analytical one, then a medium-length transition. AI models tend to generate sentences of similar length, resulting in monotonous, uniform rhythm.

How to read it: Technically, burstiness is the coefficient of variation (standard deviation / mean) of sentence word counts, normalised to a 0–1 scale. A burstiness of 0% means all sentences are the same length (AI-like). A burstiness of 60%+ indicates natural human variation. IMPORTANT: low burstiness in some sections is NORMAL, not a defect. Sections built on a fixed structural template — Abstract, Research Objectives, Hypotheses, References, bulleted lists — are inherently low-burstiness even when written entirely by a human author. The metric

measures sentence-length variance, and a 4-bullet list of research objectives shaped "To investigate… / To identify… / To evaluate…" will score around 25–40% by definition. Interpret the score in context of the section type; analytical sections (Discussion, Theoretical Background) are the meaningful places to look for low burstiness as an AI signal.

Perplexity Proxy

What it measures: True perplexity requires running the text through a language model to measure how "surprised" it is by each word. Since we avoid loading heavy ML models, this metric approximates perplexity using measurable proxies: vocabulary richness (how many unique words relative to total words), long-word usage, and sentence structure variety.

How to read it: Higher values indicate more diverse, less predictable text — characteristic of human writing. Lower values suggest formulaic, predictable patterns typical of AI output. This is a statistical approximation, not a true perplexity measurement.

Vocabulary Richness (Root TTR)

What it measures: The root type-token ratio counts unique words divided by the square root of total words. This corrects for document length — longer texts naturally repeat words more. AI-generated text tends to reuse the same vocabulary repeatedly, while human writers draw from a broader lexicon with more varied word choices.

How to read it: Values are reported as a raw score. Higher values indicate richer, more varied vocabulary. This metric feeds into the Perplexity Proxy composite score.

Limitations

- The AI detection score depends on the chatgpt-detector-roberta model, which may not cover all languages equally well. Results are most reliable for English text.
- Burstiness and perplexity proxy are statistical approximations. Short documents (< 500 words) produce less reliable scores.
- Detection technology evolves rapidly. Scores reflect the state of detection methods at the time of generation.
- Academic writing naturally shares some patterns with AI output (formal register, structured arguments). This can inflate scores.

Section-by-Section Analysis

The following table shows detection scores for each major section of the document.

Section	AI Score (Before)	AI Score (After)	Burstiness (Before)	Burstiness (After)
Abstract	N/A	N/A	94%	94%
Introduction	N/A	N/A	68%	65%
Core Concepts and Definitions	N/A	N/A	60%	60%
Review of Prior Research	N/A	N/A	70%	70%
Synthesis and Research Gap	N/A	N/A	80%	80%
Research Questions	N/A	N/A	22%	20%
Research Hypotheses	N/A	N/A	67%	64%
Research Design	N/A	N/A	75%	73%
Population and Sample	N/A	N/A	100%	100%
Research Instruments	N/A	N/A	100%	92%
Procedure	N/A	N/A	73%	73%
Data Analysis	N/A	N/A	100%	100%
Ethical Considerations	N/A	N/A	73%	73%
Characteristics	N/A	N/A	100%	100%
Descriptive Statistics	N/A	N/A	100%	100%
Hypothesis Testing	N/A	N/A	100%	100%

Section	AI Score (Before)	AI Score (After)	Burstiness (Before)	Burstiness (After)
Interpretation of key findings	N/A	N/A	74%	72%
Comparison with the literature	N/A	N/A	64%	65%
Contribution and practical implications	N/A	N/A	68%	68%
Limitations	N/A	N/A	85%	84%
Future Research	N/A	N/A	87%	87%
Summary of Main Findings	N/A	N/A	80%	78%
Conclusions	N/A	N/A	100%	100%
Recommendations	N/A	N/A	51%	51%
References	N/A	N/A	100%	100%

Note: Burstiness is a measure of sentence-length variance. Some section types are inherently low-burstiness — Abstract, Research Objectives, Hypotheses, References, and any bulleted list have fixed structural templates, so a 25–40% score for these sections reflects the format, not AI authorship. The metric is most informative on analytical sections (Discussion, Theoretical Background) where a human writer would naturally vary sentence rhythm.