

Tracing the Development of Student Research Practices in Environmental Field Science (2005–2025)

Matthew Gisi (he/him) | Graduate Student | Creative Technologies Program, Illinois State University

Faculty Mentor: Kristin Carlson, PhD | Associate Professor, Creative Technologies Program | Illinois State University

ABSTRACT

This project examines how student research practices evolve and refine over time within a long-running high school environmental field science program. The project assesses the epistemic approaches taken and quality of written reports but does not assess the validity of student research. The motivation is to measure aspects of learning considered assumed but not traced longitudinally, particularly in scientific contexts. Drawing on a subset of 20 years of digitized works from students, this archive is assessed via computational document analysis to highlight patterns in writing complexity, document structure, methodological signaling, and reflection across years. This project discusses emerging patterns that can inform potential course changes directed at the efficacy of instruction.

INTRODUCTION

High school research courses are often positioned as a means of strengthening the next generation of researchers, yet the writing they produce is rarely examined as a longitudinal record of how scientific inquiry is learned and expressed. This project begins from the premise that archives of student research do not simply document isolated epistemological development, but also reveal broader patterns in how inquiry is represented in writing, understood, and enacted over time.

It hypothesizes that discernible trends emerge in students' written research practices, reflected through structural and qualitative markers such as methods specificity, data use, and reasoning. Further, it anticipates that these markers will correlate positively with student experience and program maturity, while showing decline or disruption in response to external events, including the COVID-19 pandemic.

OBJECTIVES

1. Examine longitudinal trends in research practice markers across approximately twenty years of student reports.
2. Compare indicators of research quality across pre-COVID (≤ 2019), COVID-era (2020–2021), and post-COVID (2022–2025) cohorts.
3. Develop a structured assessment framework to inform future course adjustments, curriculum development, and instructional design.

METHOD

Corpus Assembly

Student reports were digitized from program archives and converted to plain text, a process that preserves content while flattening structural cues present in the originals. After cleaning and deduplication, 612 usable reports remained (335 dated 2005–2025, 218 undated).

AI-Assisted Scoring (LLM Pipeline)

Each report was analyzed using a locally hosted LLM (Llama 3.1 8B Instruct) configured for controlled, low-variance inference. Reliability is constructed through constraint, local inference, low-temperature configuration, structured JSON output, and multi-run consensus scoring across 2–10 independent runs per report. Reports were scored on five dimensions: writing complexity, methodological specificity, data use, reflective reasoning, and scientific reasoning (1–5 rubric).

Structural Coding

The presence or absence, not the quality, of eight canonical report components was detected through computational text analysis, allowing structure to be tracked independently of evaluative scoring.

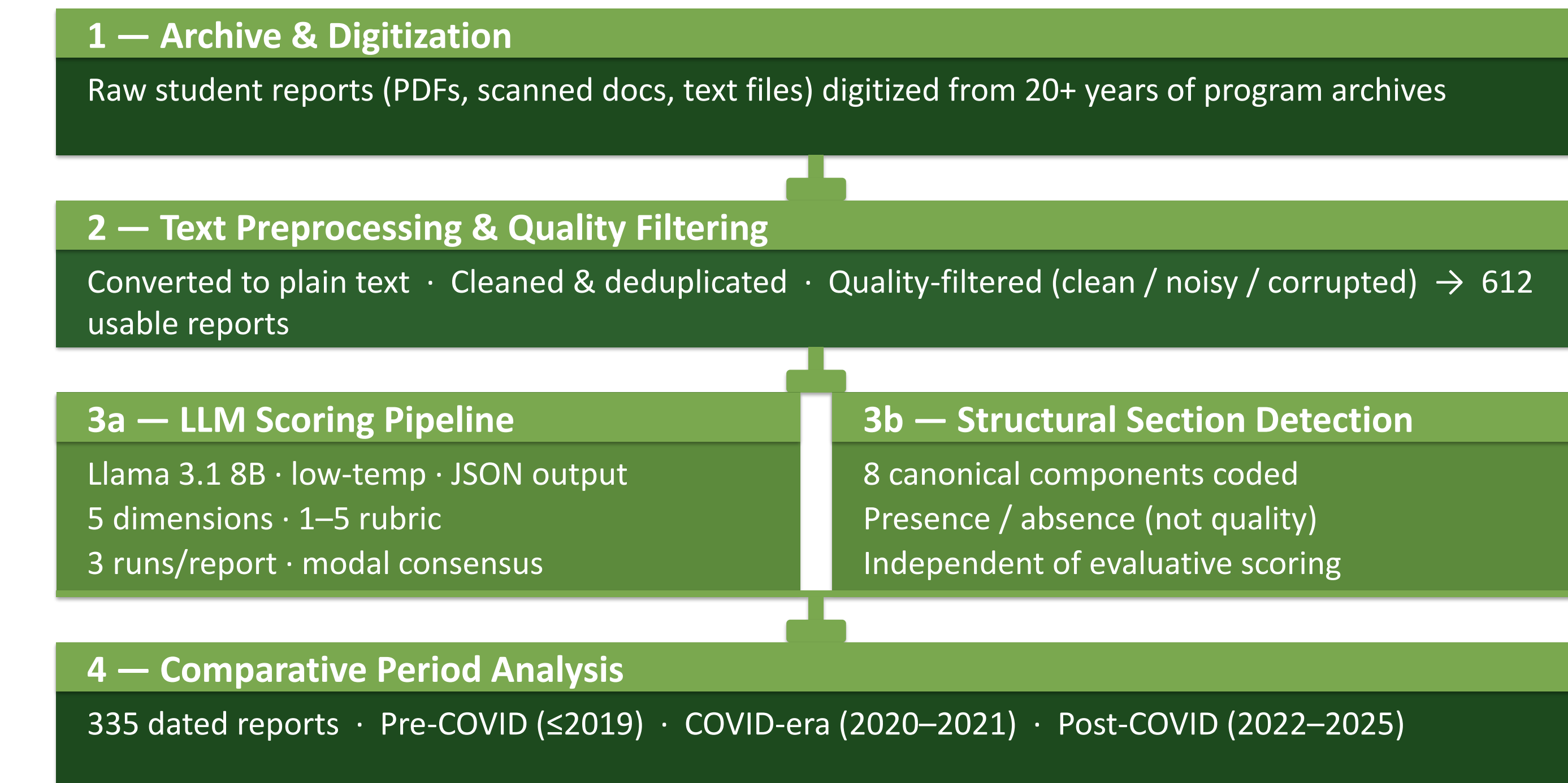
Period Analysis

Longitudinal trends were examined across the 335 dated reports. What emerges is not only change over time, but the possibility that disruption renders certain dimensions of practice more visible precisely because they fail to hold constant.

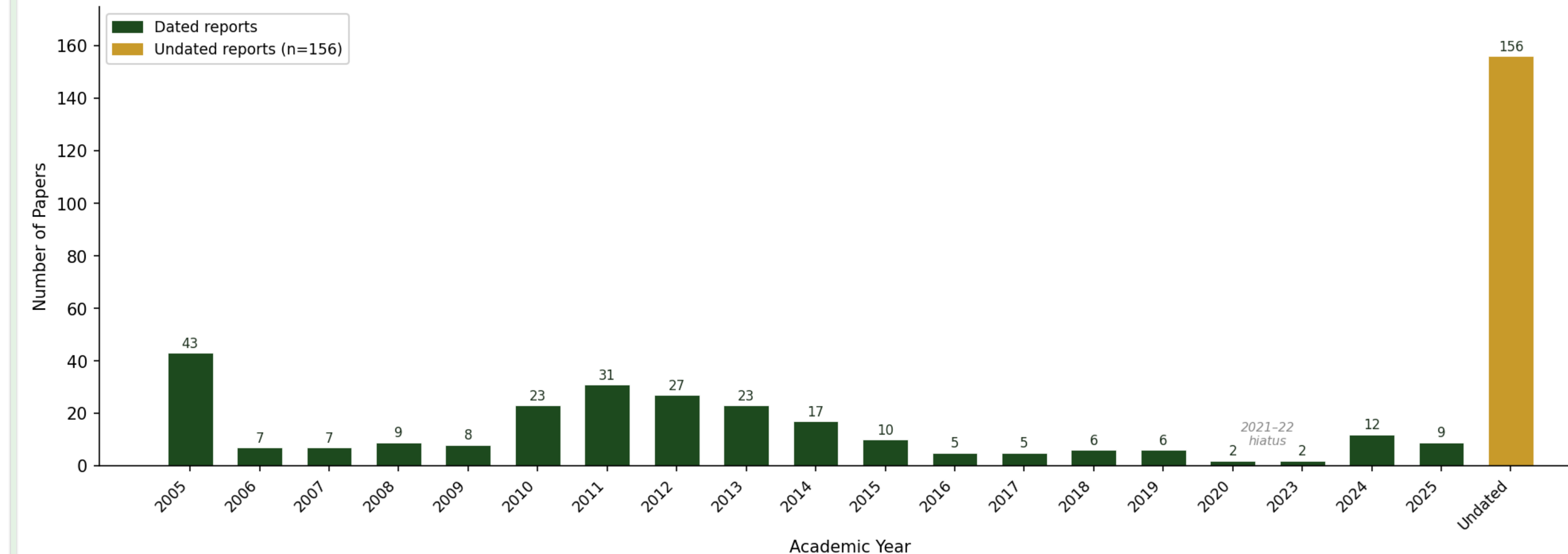
CORPUS OVERVIEW



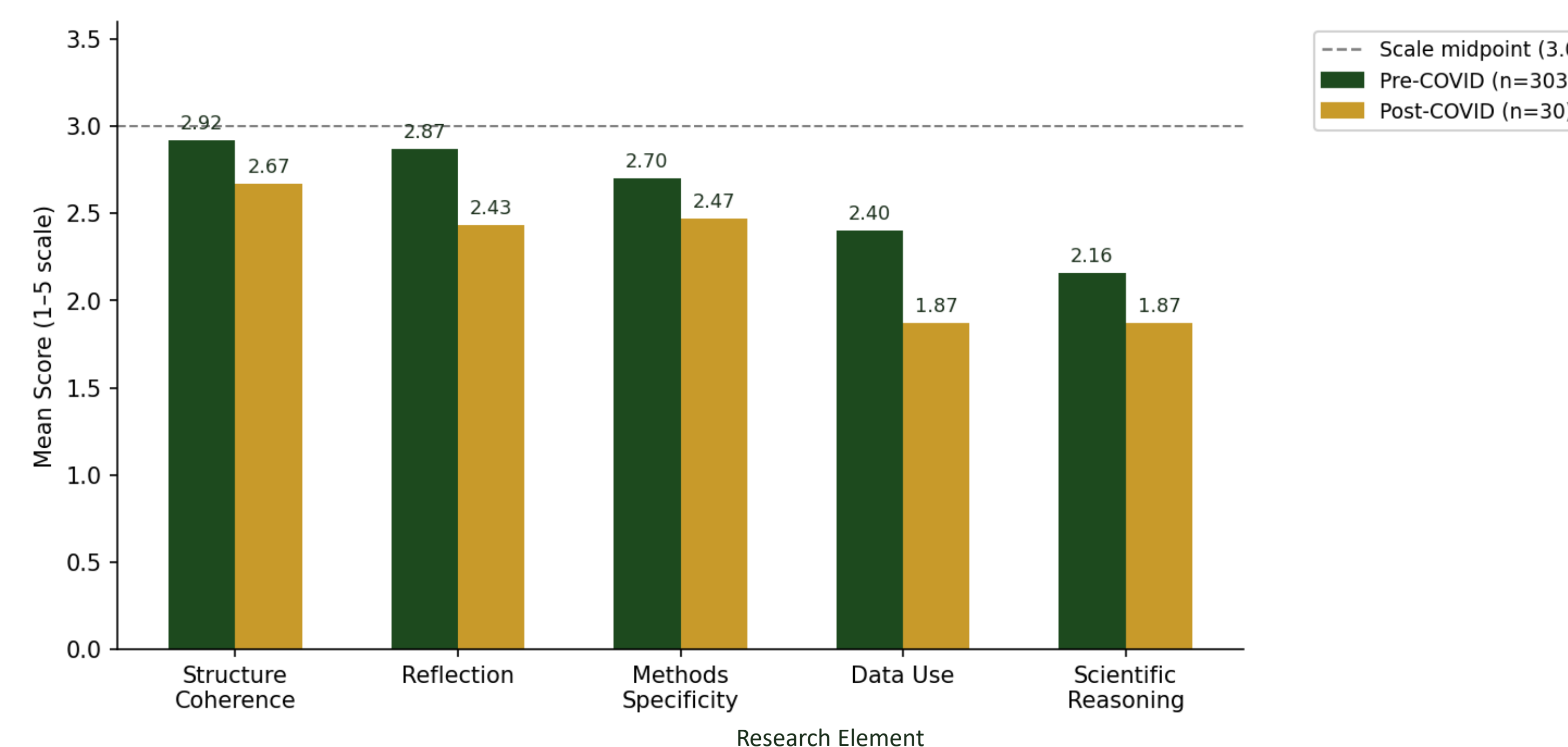
DATA ANALYSIS WORKFLOW



REPORTS PER YEAR (2005–2025 DATED CORPUS)



RESEARCH PRACTICE SCORES — MEAN (SCALE 1–5)



PERIOD COMPARISON — REFLECTION SCORE

Dimension	Pre-COVID (n=303)	Post-COVID (n=30)
Methods Specificity	2.70	2.47
Data Use	2.40	1.87
Reflection	2.87	2.43
Sci. Reasoning	2.16	1.87
Structure Coherence	2.92	2.67

RESULTS

- ▶ Methods sections are most consistently present (93%), but AI-scored quality suggests predominantly surface-level procedural reporting.
- ▶ Introduction and Conclusion sections appear in fewer than half of all reports, a structural gap that persists across the full 20-year period.
- ▶ All five research practice dimensions score below the midpoint (avg. 2.12–2.90 / 5), indicating consistent room for instructional growth.
- ▶ Post-COVID cohort shows lower scores across all dimensions compared to the pre-COVID cohort. Interpret with caution given small sample size (n=30).
- ▶ COVID-era cohort is n=2. These are drafts of projects that would have been completed summer 2020 had the COVID-19 pandemic not occurred.

CONCLUSION

Analysis throughout this project revealed structural patterns in scientific reasoning development amongst reports from a 20-year corpus. The findings of this project support a focused intervention towards complete research report structures and building stronger reflection practices.

This methodology has led to the development of a tool allowing consistent program evaluation that does not require manual scoring of hundreds of student reports. This system is also replicable in future years to provide further data down the line to assess how any structural changes from this research affect the development of student research.

FUTURE DIRECTIONS

- ▶ Extend analysis to full undated corpus by comparing student records to properly date reports.
- ▶ Examine author-specific growth over time.
- ▶ Validate AI-assisted scoring against expert human rater benchmarks.
- ▶ Further customize the model to train on a more specific scientific research context.
- ▶ Create more targeted assessment variables to identify additional elements for curriculum focus.