THE LONDON SCHOOL OF ECONOMICS AND POLITICAL SCIENCE

Analytical Report:

AI Use in Assessment in Management Courses

Linh Hoang, Jonathan Schulte, and Nayat Horozoglu

Eden Evaluation Team

© LSE 2025

Content

1. Executive Summary

2. Declaration of AI Use

3. Overview of AI Tools

4. AI Use and Grade Awards

Executive Summary

During the 2023/2024 academic year, Assistant Professor Nayat Horozoglu conducted a survey on AI use in assessments for two management courses at LSE: **MGT104 Operations Management** and **MGT308 Simulations for Managerial Decisions**.

All students were asked to declare their use of AI in the completion of the summative assessment following the courses' AI policy. There were two rounds of surveys to ensure all students declared their use of AI in assessment.

Based on the analysis conducted by the Eden Evaluation Team, the findings revealed a growing prevalence of AI-supported practices in assessment, identified the most commonly used tools, and illustrated patterns of AI use across different learning purposes. **Key findings include**:

- Ultimately, 42% of students in MG104 and 34% of students in MG308 declared using AI for the assessment;
- There is a wider variety of AI tools used in MG104 compared to MG308, though ChatGPT and Grammarly remain to be the most popular in both courses;
- AI tools are majorly used for (1) literature searching / idea generation, (2) assessing sentence and wording / proofreading, and (3) checking grammar / spelling / punctuation in both courses.

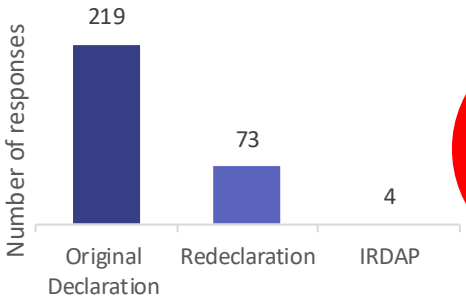
The study also investigated whether there is an association between AI use and grade awards. Descriptive statistical analysis showcased that **grades are not statistically significantly different with or without the use of AI**. Furthermore, **students using AI have more consistent grades** in both courses, while non-AI users' grade distribution shows greater variability, indicating a mix of strong and weak performances.

Declaration of AI Use

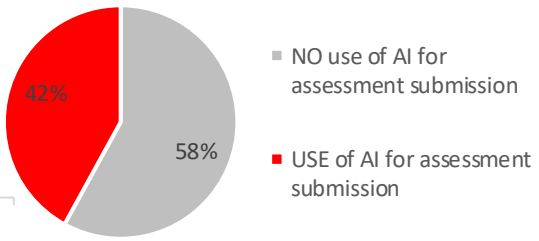
This section highlights the key findings on the two rounds of declarations of AI use in assessment in two LSE courses: MG104 and MG308.

MG104

Types of Declaration of AI Usage

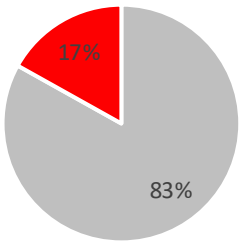


Summary of Two Declarations



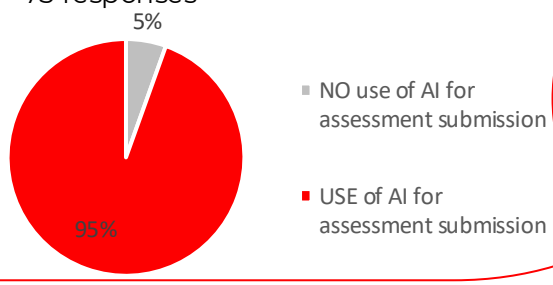
Original Declaration

219 responses



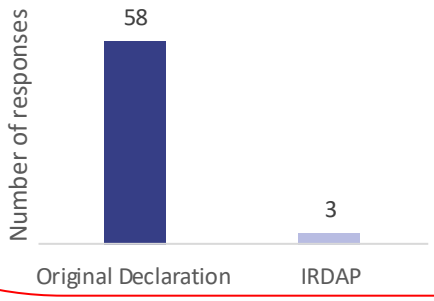
Redeclaration

73 responses



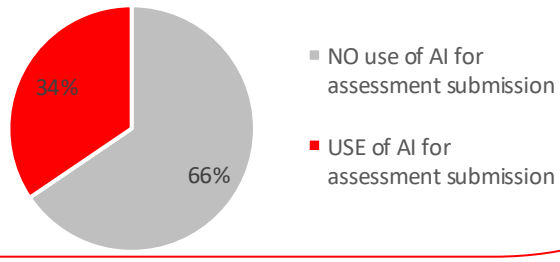
MG308

Types of Declaration of AI Usage



Original Declaration

58 responses



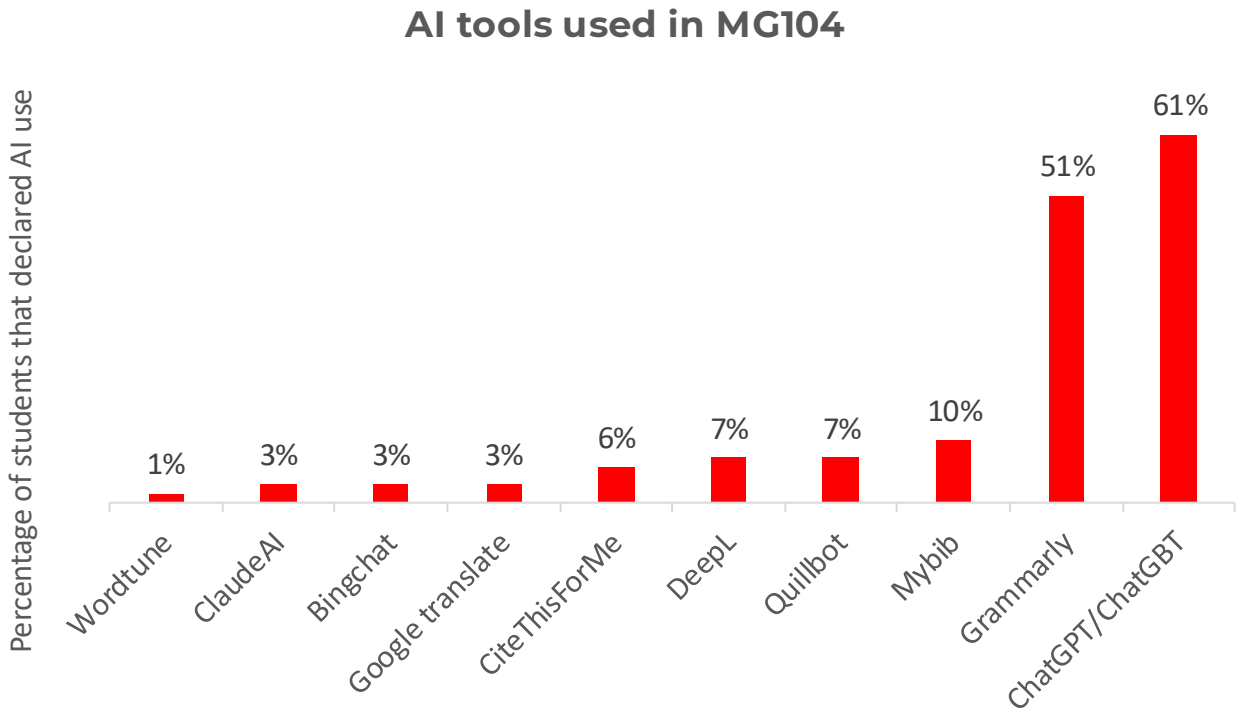
Comment

- Among 182 students in MG104 that originally declared no use of AI, 55 changed their declaration from no to yes in the redeclaration survey, increasing the percentage of declaration from 17% to 42% ultimately.
- The percentage of students using AI tools is significantly higher for MG104 compared to MG308.

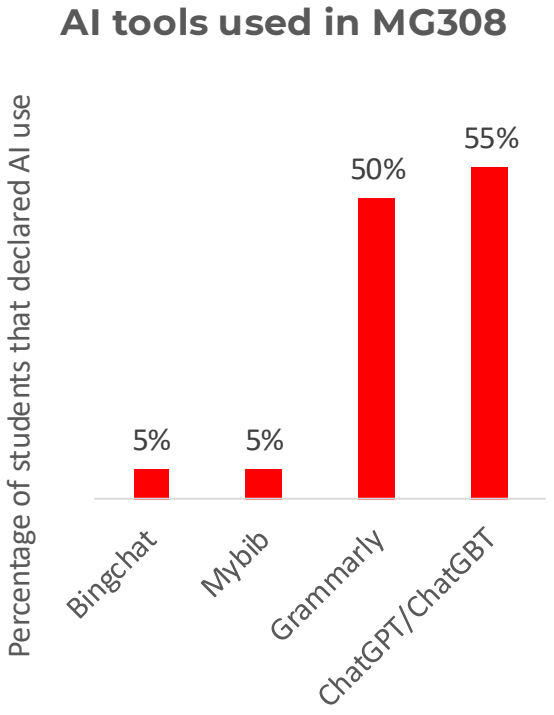
Overview of AI Tools

What are the AI tools used by students?

The following graphs present **the percentage of students that declared AI use** by each tool in the **redeclaration for MG104** and **original declaration for MG308** (one student can use multiple tools).



There is **a wider variety** of AI tools used in **MG104** compared to MG308, though **ChatGPT and Grammarly** remain to be the most popular in **both courses**.



Overview of AI Tools

What did students use AI tools for?

In both courses, AI tools are majorly used for (1) **literature searching / idea generation**, (2) **assessing sentence and wording / proofreading**, and (3) **checking grammar / spelling / punctuation**. These codes are generated inductively based on survey responses.

MG104 Redeclaration	Tools	Literature searching / Idea generation	Overall structure	Sentence and Wording / Proofreading	Grammar check / Spelling / Punctuation	Translation	Referencing / Bibliography / Citations	Code fixing
No. of times used	ChatGPT/GBT	23	8	35	14	2	2	
	ClaudeAI	2		1	1			
	Bingchat	3						
	DeepL			2		4		
	Grammarly		5	10	53	1		
	Quillbot			3	3	1	1	
	Wordtune			1	1			
	Google translate					2		
	Mybib			2			8	
	CiteThisForMe						4	

In **MG104 (Redeclaration)**, **ChatGPT** is preferred for assessing sentence and wording / proofreading, while **Grammarly** is frequently used for checking grammar / spelling / punctuation. Additionally, **Mybib** is the preferred tool for generating reference / bibliography / citations.

MG308 Original Declaration	Tools	Literature searching	Overall structure	Sentence and Wording / Proofreading	Grammar check / Spelling / Punctuation	Translation	Referencing / Bibliography / Citations	Code fixing
No. of times used	ChatGPT/GBT	4		10	5		1	2
	Bingchat	1						
	Grammarly		2	8	11			
	Mybib						1	

In **MG308 (Original Redeclaration)**, **ChatGPT** is also preferred for assessing sentence and wording / proofreading, while **Grammarly** is also frequently used for the same purpose in addition to checking grammar / spelling / punctuation. Noticeably, **ChatGPT** is particularly used for **solving coding issues** among students in this course.

LSE THE LONDON SCHOOL OF ECONOMICS AND POLITICAL SCIENCE

AI Use and Grade Awards

This section investigates whether there is an association between AI use and and final grades in the two LSE courses: MG104 and MG308.

MG104

Descriptive statistical analysis of grades with and without AI use

Grades with AI	
Mean	60.28
Standard Error	0.81
Median	61.50
Mode	62.00
Standard Deviation	7.88
Sample Variance	62.03
Kurtosis	1.44
Skewness	-0.38
Range	48
Minimum	35
Maximum	83
Sum	5666
Count	94

Grades WITHOUT AI	
Mean	61.63
Standard Error	0.84
Median	62.00
Mode	58.00
Standard Deviation	9.20
Sample Variance	84.62
Kurtosis	1.52
Skewness	-0.26
Range	56
Minimum	31
Maximum	87
Sum	7457
Count	121

Independent t-test for mean difference: p-value of 0.257 indicated **grades are not statistically significantly different** with or without the use of AI

MG308

Descriptive statistical analysis of grades with and without AI use

Grades with AI	
Mean	64
Standard Error	3
Median	64
Mode	79
Standard Deviation	15
Sample Variance	238
Kurtosis	1
Skewness	-1
Range	59
Minimum	27
Maximum	86
Sum	1273
Count	20

Grades WITHOUT AI	
Mean	59.46
Standard Error	2.49
Median	60.00
Mode	60.00
Standard Deviation	14.74
Sample Variance	217.37
Kurtosis	0.02
Skewness	-0.72
Range	63
Minimum	23
Maximum	86
Sum	2081
Count	35

Mann-Whitney test for mean difference: p-value of 0.845 indicated **grades are not statistically significantly different** with or without the use of AI

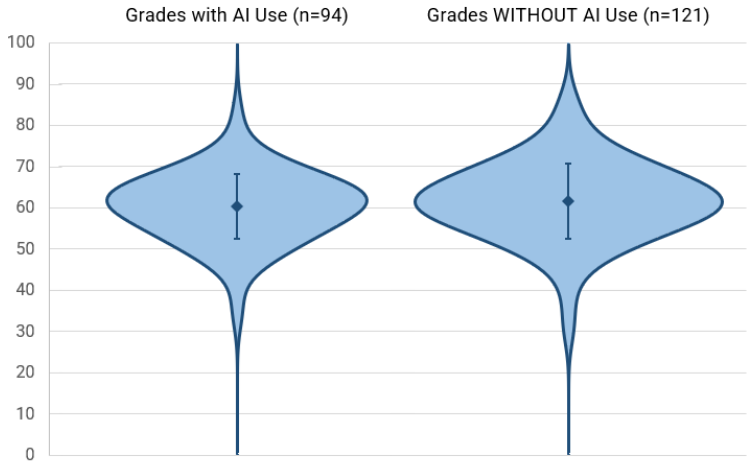
AI Use and Grade Awards

MG104

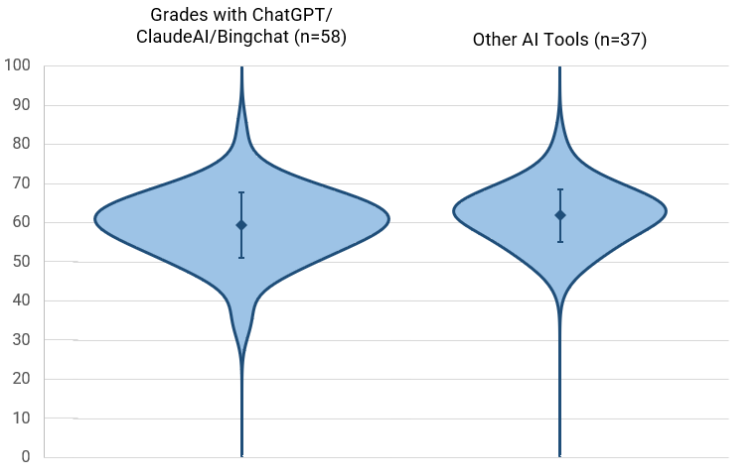
Comparing distribution of grades with and without AI use

- **Students using AI** have more **consistent grades**, with a narrower spread around 60-70. The sharp peak indicates that **most students achieve grades in the same range**, reducing outliers.

- **Non-AI users**' grades show **greater variability**, with grades distributed more widely. This group displays **more extreme outcomes**, with some students excelling and others underperforming significantly.



Comparing distribution of grades with AI use including and excluding ChatGPT/ClaudeAI/Bingchat



- The grades with AI tools **including ChatGPT/ClaudeAI/Bingchat** have **broader variability** compared to that **excluding these three**, with significantly **more underperforming extremes**.

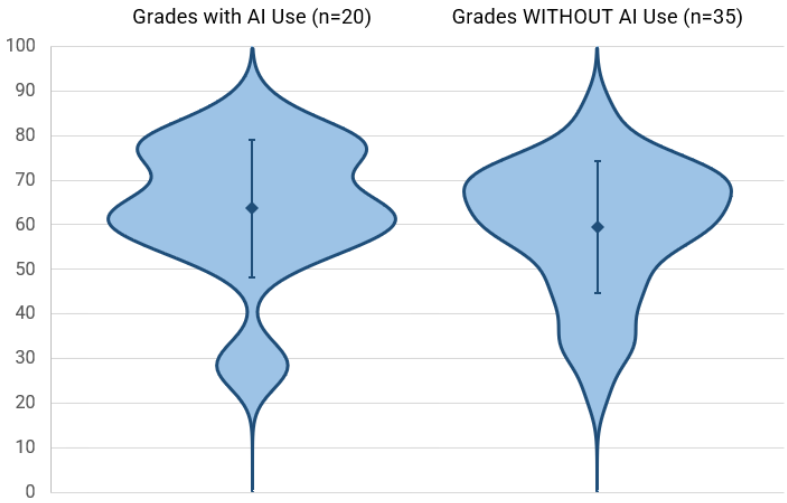
- The **former**'s distribution is spread out around the central range of **50-60**, while the **latter**'s is concentrated around **55-65**.

AI Use and Grade Awards

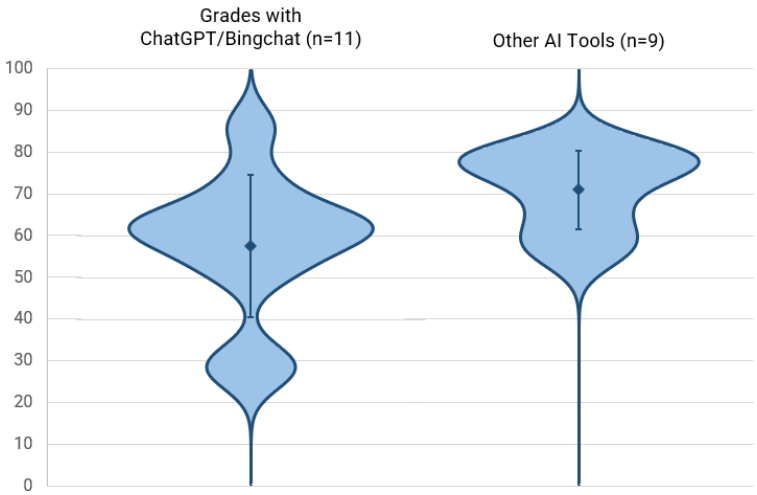
MG308

Comparing distribution of grades with and without AI use

- **Students using AI** have the more **consistent grades**, with a more defined peak around 60-80.
- There is a **noticeable dip below 40**, indicating a small subset of students scored quite low even with AI.
- **Non-AI users** exhibit **wider grade distribution**, indicating a mix of strong and weak performances.



Comparing distribution of grades with AI use including and excluding ChatGPT/Bingchat



- The grades with AI tools **including ChatGPT/Bingchat** have broader variability compared to that **excluding these two**, with **two noticeable peaks** (around 60-70 and around 30-40), and **a dip below 40**.
- **The use of AI tools excluding ChatGPT/Bingchat** has more **uniform** grades clustered around 60-80, with narrower tails, suggesting fewer extreme low or high grades.