Valko *et al. International Journal for Educational Integrity* (2026) 22:2
https://doi.org/10.1007/s40979-026-00212-2

International Journal for Educational Integrity

RESEARCH Open Access

Unmasking CAMEO cheating in MOOCs via behavioral and temporal analysis without IP tracking

Matúš Valko[1*], Michal Munk[1,2] and Dasa Munkova[1]

*Correspondence:
Matúš Valko
matus.valko@ukf.sk
[1]Department of Informatics, Constantine the Philosopher University in Nitra, Nitra, Slovakia
[2]Science and Research Centre, University of Pardubice, Pardubice, Czech Republic

Abstract

The rapid growth of online education has introduced new challenges in maintaining academic integrity, particularly in MOOCs where sophisticated cheating strategies, such as CAMEO (Copying Answers using Multiple Existences Online), have emerged. This study proposed a novel method to detect CAMEO-style cheating by identifying suspicious harvester and master accounts without relying on IP address tracking, which can be unreliable in shared or masked network environments. Using behavioral analytics and temporal patterns of task engagement, we analyzed data from 558 accounts in a Java programming course on the Priscilla MOOC platform. Our findings indicated that harvester accounts exhibited minimal engagement, low variability in task repetitions, answer purchases, and time spent on tasks, but high variability in relative scores. Immediate-mode CAMEO behavior was detected effectively, while batch-mode behavior remained more challenging due to delayed submissions. The method highlighted the importance of behavioral variability metrics in distinguishing suspicious accounts, providing an alternative to traditional IP-based detection methods. These results underscore the potential for behavioral analytics to strengthen academic integrity in online learning environments, while emphasizing the need for further research to validate and generalize detection methods across larger and more diverse datasets.

Clinical trial number

Not applicable.

Keywords MOOC, CAMEO, Cheating, Programming course, Variability

Introduction

In today's rapidly evolving digital landscape, organizations face increasingly complex challenges in combating online fraud, which threatens both financial assets and reputational integrity (Sorour et al. 2024). Implementing robust fraud detection strategies is therefore essential for effectively mitigating these threats. Moreover, this urgency is heightened by the fact that the world of online fraud is constantly evolving, as fraudsters constantly derive new strategies to deceive systems and exploit weaknesses in

organizational processes. As a results, and in response to this persistent innovation in fraudulent methods, organizations must not only react to known threats but also remain proactive by implementing strong fraud prevention measures before they occur (Wu et al. 2024).

To achieve this, one of the most effective approaches involves detecting deviations from expected patterns by monitoring and analyzing users' behavior. In this context, behavioral analytics plays a crucial role by enabling the identification of suspicious activity, including irregular login locations, sudden shifts and changes in engagement patterns, or unusual transaction amounts. Additionally, by leveraging these insights, organizations gain a deeper understanding of user behavior patterns, which turn enhances the accuracy and precision of anomaly detection. Ultimately, the strategic integration of behavioral analytics strengthens fraud detection framework, allowing organizations to identify threats earlier with greater reliability. Similarly, this principle of early detection and pattern-based monitoring extends beyond financial domains and into educational environments. In an academic context, fraud is defined as dishonest or deceptive behavior intended to gain an unfair advantage, including the misuse of data, resources, equipment, or methods to complete academic tasks in a dishonest manner (Salamah 2022). It also refers to actions or behaviors undertaken by a student (such as plagiarism, cheating, fabrication, and contract cheating) that are intended to deceive for unfair advantage by violating academic regulations (Eckstein 2003; Tertiary Education Quality and Standards Agency 2025). In recent years, academic cheating has become an increasing concern across educational environments, particularly in schools, higher education institutions, and universities. Accordingly, several studies have sought to investigate the causes of academic dishonesty and propose effective countermeasures (Aruğaslan 2024; Colnerud and Rosander 2009; Crittenden et al. 2009; Zhao et al. 2022). To mitigate these risks, universities must secure online examination platforms, enforce explicit academic integrity policies, and cultivate ethical artificial intelligence (AI) practices alongside digital literacy among students (Jenkins et al. 2023; Sozon et al. 2024).

In response to these challenges, several preventative and detection mechanisms have been employed in online learning and assessment environments, such as randomized question selection from question pools, time restrictions, questions and answers randomization, browser tab locking, user authentication, and online proctoring (Taşkın 2024). These strategies are increasingly supported by machine learning techniques and AI (Salih et al. 2025; Wang et al. 2024), rule-based algorithms (Northcutt et al. 2016), and behavioral analytics (Yuan et al. 2024). Among these approaches, our study specifically emphasizes behavioral analytics due its capacity to reveal real-time irregularities in user engagement. Prior research indicates that 42% to 97% college students admitted to having engaged in some form of cheating, depending on demographic and geographic factors (Erguvan 2022; Hasri et al. 2022; Lingán-Huamán et al. 2024; Yu et al. 2023).

Despite these mitigation efforts, however, detecting and preventing cheating in online assessments continues to be one of the most serious obstacles and remains one of the biggest challenges facing remote assessment systems. Traditionally, social stigma has served as a key behavioral deterrent, but this influence weakens when students have no negative attitude towards cheating, increasing the likelihood to engage in such practices (Noorbehbahani et al. 2022; Salehi and Gholampour 2021). For this reason, analyzing student behavioral engagement is critical, not only for performance prediction but also

for abnormality detection. As previous studies demonstrate, behavioral engagement correlates strongly with academic achievement and can serve as a valuable predictor of overall academic outcomes (Taşkın and Kokoç 2025). Typical indicators include metrics from learning management systems (LMSs), active participation in synchronous courses, frequency of LMS access, time spent on tasks, and data on course access or assignment completion (Caspari-Sadeghi 2022).

Building on this foundation, our research investigates cheating behavior patterns in asynchronously submitted assignments on the micro-learning platform Priscilla (Skalka & Drlik, 2018). In doing so, we propose a method inspired by Northcutt et al. (2016) to identify suspicious "harvester" accounts, where students create secondary accounts to extract correct answers and later submit them through a primary "master" account.

The objective of this research is twofold: (1) to detect students employing the Copying Answers using Multiple Existences Online (CAMEO) cheating method, and (2) to introduce a reliable fraud detection approach in the Priscilla Massive Open Online Courses (MOOCs) platform. Furthermore, we analyze the behavioral engagement patterns associated with the identified CAMEO accounts, placing particular emphasis on behavior variability as a marker of account abnormality. These contributions aim to expand current understanding of cheating within MOOCs and strengthen anomaly-driven fraud detection models in educational settings.

The rest of the article is structured as follows. The next section outlines related work on academic dishonesty, synthesizing prior studies that address similar detection challenges. Following that, the Methodology section describes our proposed approach, dataset, and hypotheses for identifying CAMEO users. Subsequently, the Results section presents key findings, which are then interpreted in the Discussion section to explore emerging cheating behaviors and their impact on MOOCs. Finally, the Conclusion section summarizes our contributions, while acknowledging the study's limitations and implication for education.

Related work

Research consistently demonstrates that behavioral engagement indicators, particularly students' interaction with course materials, offer valuable insights into academic performance and can also reinforce the integrity of online assessment processes (Taşkın and Kokoç 2025). Completing this perspective, Odongo et al. (2021) further showed that students' perception of cheating as a "victimless crime", combined with insufficient institutional mechanisms, contributes to conditions that normalize academic dishonesty.

From a methodological perspective, Du et al. (2022) classified cheating detection methods in online assessments into two broad categories: (1) answer similarity and (2) time-based behavioral analysis. Specifically, answer similarity methods rely on pairwise comparisons of student responses and remain among the most widely adopted approaches. These methods frequently employ measures such as text similarity metrics or plagiarism detection tools like Measure of Software Similarity (MOSS) and its extended version Temporal Measure of Software Similarity (TMOSS), particularly in programming-based online courses. Conversely, time-based analysis focuses on irregularities in submission behavior, including outliers, abnormally rapid completion, or suspicious synchronization across student response timelines. Within this approach,

Ranger et al. (2020) proposed measuring outliers using response time, revision time, and speed-accuracy relation.

One of the earliest timeline-driven MOOC cheating detection frameworks was introduced by Northcutt et al. (2016), specifically designed to identify students using the multiple -account strategy known as CAMEO. In this method, a student creates one or more "harvester" accounts to obtain correct answers, which are then used by their main "master" account to complete assessments. Accordingly, comparing account interaction timelines became a core heuristic for identifying cheating-oriented duplicate accounts (harvester accounts), an approach that directly informed the method adopted in our research.

However, as the use of online learning platforms increases, additional challenges have emerged, such as advanced multi-account cheating behaviors (CUMA, Copying Using Multiple Accounts). In practice, a CAMEO user typically operates at least one harvester account, retrieving correct answers either through systematic guessing or by exploiting built-in platforms feedback mechanisms, including the instructor-enabled "Show Answer" option (Bao et al. 2017; Northcutt et al. 2016). In this context, Alexandron et al. (2015) proposed an IP-based algorithm that detected 49 fraudulent certificate earners (representing 9.8% of all accounts) who had used harvester accounts at least once. Their analysis revealed two distinct behavioral patterns among CAMEO users (Fig. 1):

- Instant (immediate) mode - a user retrieves the correct answer using a harvester account and submits it via the master account within a short time interval. Harvesting and answer submission occur in rapid cycles, frequently within 5–15 min (Alexandron et al. 2015; Northcutt et al. 2016).
- Batch mode - this pattern involves using the harvester account to obtain correct answers to multiple questions, which are then submitted by the master account. Answers are pre-collected and submitted at high speed (e.g., every 20–30 s), often in sequences that may not match the original harvesting order. Some studies also

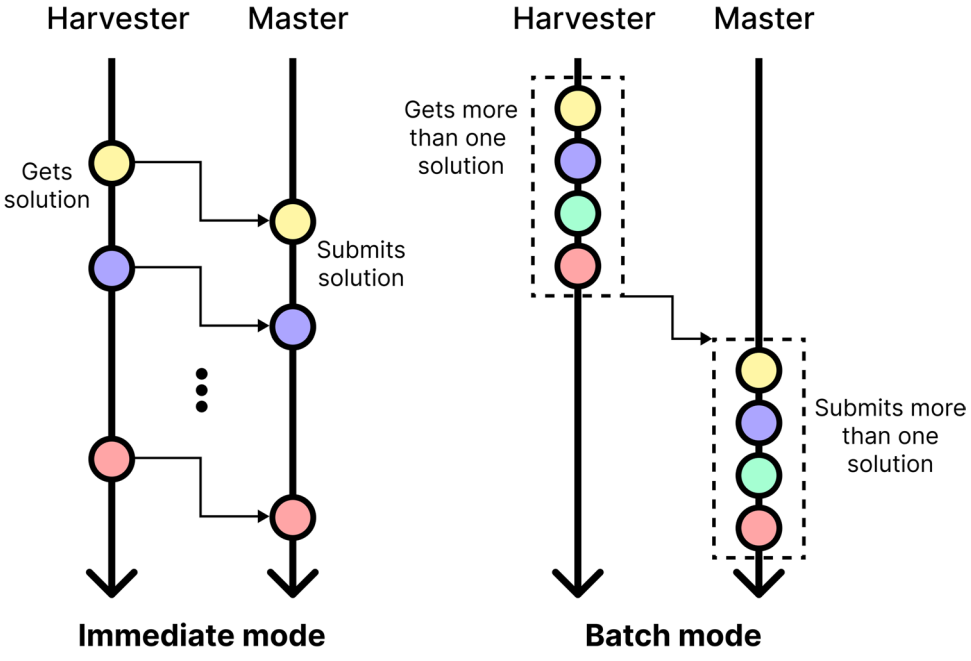

Fig. 1 Two behavioral patterns of CAMEO users

applied a threshold of 10 rapid-correct submissions (10 questions) for flagging suspicious behavior (Alexandron et al. 2015).

Both modes (behavioral patterns) may combine into hybrid or alternating patterns, further increasing detection complexity. Nevertheless, immediate mode is generally easier to detect due to its short and repetitive harvesting-submission cycles, conceptually resembling "close submitter" behavior. Despite this similarity, however, the two are not equivalent: CAMEO detection emphasizes newly created, harvesting-dominant accounts, whereas close submitter detection focuses on regular enrolled student accounts that collaborate excessively during coursework.

Jaramillo-Morillo et al. (2020) identified academically suspicious coordination using both answer similarity and temporal proximity, applying a 5-minute threshold., although they did not detect CAMEO behavior due to platform restriction (new accounts creation was prohibited). Likewise, Balderas et al. (2021) and Balderas and Caballero-Hernández (2020) showed sequential test-taking collaboration, detected through grade similarity and aligned submission timestamps.

Additionally, Ruipérez-Valiente et al. (2017) found that such "close submitters" exhibited statistically lower engagement with MOOC learning resources, suggesting systematic collaboration or dishonest practices, which raises concerns about certificate credibility. Supporting this argument, Northcutt et al. (2016) highlighted three defining characteristics that distinguish CAMEO cheating in MOOCs: (1) user self-sufficiency (CAMEO users duplicate their efforts by creating multiple accounts and copying answers from themselves; this independence increases the method's accessibility and makes it harder to detect), (2) exploitation of asynchronous access structures (CAMEO users can exploit the system by applying the strategy question-by-question, potentially completing the course in a single session; the lack of content access restrictions facilitates this behavior), and (3) unregulated enrollment mechanisms that may inadvertently enable credential fraud.

Building on these characteristics, Northcutt et al. (2016) proposed an algorithm centered on five filtering assumptions: (1) prior harvesting of correct answers (the harvester account must obtain the correct answer before the master submits it), (2) short communication gap between accounts (the harvester account should provide answers to the master account within a short time interval), (3) harvester disinterest in certification, (4) shared IP address and history, and (5) minimal overlap with other accounts sharing the same network.

Subsequently, Bao et al. (2017) reproduced the study across ten MOOCs, confirming CAMEO behavior in 9 out of 10 courses and estimating that 1.9% of certificates were likely achieved through CAMEO-driven fraud.

Alexandron et al. (2017) conducted deeper log-level behavioral analysis on the edX platform (using edX log files) and reported substantially higher cheating prevalence – 10% vs. 2.5% reported by Northcutt et al. (2016) for the same course. They attributed this to broader harvesting detection scopes (show answer, exhaustive guessing, metadata), inclusion of batch behavior, and longer suspicious response thresholds (24 h vs. 5 min). Their findings further classified CAMEO behavior as either (a) help-seeking (post-failure assistance) or (b) premeditated (strategic shortcut use), showing that 78% of certificate-linked CAMEO interactions were deliberate and pre-planned, reflecting a primarily outcome-driven rather than learning-driven intent.

Consistent with this trend toward AI-supported detection, Alexandron et al. (2016, 2019) applied a logistic regression using person-fit features (Guttman error and standard error) and learning analytics features (fraction of videos watched and fraction of correct answer in less than 30 s) to construct a universal cheating classifier. Similarly, Ruiperez-Valiente et al. (2017) employed a random forest model enhanced through VSURF, concluding that temporal features related to submission latency, video engagement time, and first-attempt accuracy offered the strongest signals for harvested answer detection.

these studies expose two recurring limitations in prior work: reliance on IP metadata and insufficient adaptation to asynchronous student interaction behavior. For this reason, and to address these constraints, our study adopts a behavioral analytics-first approach. By prioritizing engagement variability, task interaction timing, and performance consistency, we aim to develop a more reliable, scalable, and context-aware method for detecting fraudulent multi-account cheating behaviors - particularly CAMEO - in modern MOOC environments (such as Priscilla).

Methodology

According to Ranger et al. (2020), indicators of academic dishonesty can be conceptualized across multiple analytical dimensions, including: dependence on latent trait models, specificity to a particular cheating form, requirement of a priori hypotheses regarding affected questions or students, and reliance on either response content or behavioral process data (e.g., submission timestamps, revisions, completion duration). Building on this, the present study focuses specifically on detecting and linking master and harvester accounts exhibiting behavioral signs of the CAMEO cheating strategy.

To achieve this goal, our methodology is structured into series of filtering and linkage stages. First, we identify suspected harvester accounts, which are commonly created solely to extract correct responses rather than to meaningfully engage in coursework. Consequently, these accounts tend to demonstrate both minimal course interaction and low cumulative performance score. On the target platform investigated within this research, students earn 10 points for completing basic tasks (e.g., single-choice, multiple-selection, short answer) and 20 points for more complex task types, such as programming responses. Parallel to this point-based structure newly registered accounts are also provisioned with 100 platform-specific virtual currency units ("coins"), which can be exchanged to reveal correct answers via a "Buy Answer" feature. Although this feature was originally intended to support comprehension of difficult or poorly formulated questions, in practice, it can be exploited to power multi-account cheating behaviors – including answer harvesting for use in a separate master account.

Considering these operational constraints, we estimate that a harvester account using only its initial coin allocation could purchase answers for up to ten low-cost tasks (10 coins each), five high-cost tasks (20 coins each), or a hybrid combination of both. Accordingly, an account used exclusively for harvesting may complete no more than ten tasks, yielding a maximum possible course score of 100 points. For this reason, we classify accounts completing zero to ten tasks (≤ 100 points) as suspicious candidates. Following this rationale, we formally define a suspicious harvester account as one that simultaneously meets the criteria of: (a) total course score ≤ 100 points, and (b) number of tasks (answers) obtained via the "Buy answer" feature ≤ 10.

Next, after isolating suspicious harvester accounts, the second stage involves linking them to possible master accounts. A key requirement for validating such a linkage is temporal precedence – namely, the harvester account must acquire the correct answer before the master submits that same answer. However, because CAMEO usage patterns may manifest in either immediate or batch mode, the time interval between harvesting and submission may vary.

In immediate mode, in particular, harvesting–submission cycles are expected to occur within a short and repetitive time span. Thus, consistent with thresholds adapted in previous (Bao et al. 2017; Jaramillo-Morillo et al. 2020; Northcutt et al. 2016), we applied a 5-minute temporal window between harvesting and submission. Importantly, unlike Northcutt (2016), we do not rely on IP address matching when linking harvester and master accounts. This decision is justified by the fact that:

- single harvester accounts may serve multiple master accounts in coordinated cheating networks,
- institutional settings such as university campuses or dormitories often involve share network infrastructures that generate false IP overlap, and.
- students can easily mask or rotate IPs using VPNs or proxy servers, reducing the reliability of IP-based identification.

As Ruiperez-Valiente et al. (2017) argue, detection mechanisms based on behavioral rather than network metadata may offer greater accuracy and resilience in asynchronous learning environments.

Based on these observations and constraints, we identify harvester-master pairs that meet the conditions of:

- harvester account obtained the correct answer using the "buy answer" feature for a specific task, and.
- master account submitted the correct answer to the same task within five minutes of the harvester's purchase and received the task's full possible score.

Finally, after extracting all temporally valid master-harvester candidate pairs, apply an additional refinement filter by retaining only master accounts linked to more than 10 harvester accounts. By doing so, we reduce the probability of accidental flagging caused by isolated matching submission times or rare non-CAMEO behaviors. This step aligns with prior findings from Alexandron et al. (2017), who emphasize that authentic master-harvester cheating profiles are typically characterized by high-frequency, repeated interactions across many assessment tasks (questions) rather than sporadic or low-scale coincidences.

Dataset

To conduct our analysis, we examined user interaction behavior within the Priscilla MOOC platform (priscilla.fitped.eu). Given the platform's multidisciplinary course repository, we narrowed our focus to the mandatory first-year Java programming course from the in-person Applied Informatics bachelor's program, which incorporates asynchronous micro-learning through Priscilla. However, although the platform follows an open-registration model, external users have minimal incentive to cheat, as Priscilla does not grant certificates or academic credit beyond access to learning content.

Accordingly, the dataset consisted of 558 registered student accounts that interacted with the Java programming course between January 2021 and December 2023. A notable affordance of this platform is the "Buy Answer" feature, which allows users to reveal correct solutions using virtual currency. While intended to support comprehension of unclear or cognitively demanding questions, this mechanism can be systematically misused in multi-account strategies such as CAMEO, where harvester accounts extract answers that are later submitted by a master account to bypass course requirements.

To support robust and noise-controlled analysis, we used two primary log datasets:

- *x_logs* (1,922,649 records): general behavioral interaction logs, capturing actions such as open/ close tasks. In particular, we analyzed "Buy Answer" action (feature). Accounts that achieved low overall course performance (≤ 50 points, equivalent to 0–5 completed tasks) and used "Buy Answer" feature at least once were initially labeled as suspicious harvester candidates.
- *user_attempts* (298,044 records): detailed submission logs capturing task identifiers, scores, and timestamps. These data enabled both (a) verification of the suspected harvester accounts and (b) temporal and performance-based linkage with master accounts.

Subsequently, applying the filtering logic across both datasets revealed 38 potential harvester accounts and 60 tasks exhibiting evidence of CAMEO-style cheating. Building on this filtered task set, we generated a transformed, analysis-focused subset of 377 accounts that had attempted one or more of the 60 suspicious tasks. By doing so, we reduced behavioral noise, minimized false positives, and ensured that anomaly detection remained constrained to shared assessment contexts.

The transformed dataset contains the following account characteristics (Fig. 2):

- A_ID - unique account identifier,
- RS (Relative Score) - ratio of obtained to maximum possible points,
- R (Repetitions) - total number of task attempts,
- BA (Buy answer) - frequency of answer purchases,
- T (Time) - total time spent on tasks.

These characteristics (RS, R, BA, and T) were computed for all course tasks (TOTAL, green background in Fig. 2) and separately for only the 60 suspected tasks (SAMPLE, yellow background in Fig. 2), enabling comparative behavioral profiling across coursework scope.

Additionally, three task types were behaviorally aggregated per account, producing 12 summarized characteristics (3 task types × 4 measures = 12 columns, grey background in Fig. 2). For each account and task type, we computed:

- R (Repetitions) - number of attempts per task,

A_ID	RS_TOTAL	R_TOTAL	BA_TOTAL	T_TOTAL	RS_SAMPLE	R_SAMPLE	BA_SAMPLE	T_SAMPLE	R_8	SAMAX_8	BA_8	T_8	R_5	SAMAX_5	BA_5	T_5	R_2	SAMAX_2	BA_2	T_2	SA
145	0.98	115	0	19479	1.0	18	0	5740	0.346	0.096	0.0	110.385	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0
540	0.91	294	0	58399	0.97	43	0	21424	0.808	0.173	0.0	411.154	0.25	0.25	0.0	11.0	0.0	0.0	0.0	0.0	0
560	1.0	238	17	47852	1.0	81	7	23283	1.558	0.327	0.135	447.75	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0
566	0.95	126	0	9642	1.0	4	0	470	0.077	0.019	0.0	9.038	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0
567	0.97	403	0	70383	0.96	125	0	18753	2.385	0.558	0.0	360.212	0.0	0.0	0.0	0.0	0.25	0.25	0.0	5.5	0
568	1.0	40	0	3323	1.0	5	0	908	0.096	0.038	0.0	17.462	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0
570	0.86	174	1	10149	0.83	58	0	4440	1.115	0.231	0.0	85.385	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0
572	0.93	596	2	199399	0.89	114	2	56482	2.135	0.423	0.038	1082.808	0.75	0.0	0.0	44.0	0.0	0.0	0.0	0.0	0
573	0.87	142	0	4287	0.78	11	0	313	0.212	0.135	0.0	6.019	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0
574	0.99	146	4	5227	0.99	42	1	447	0.808	0.712	0.019	8.596	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0

Fig. 2 Portion of the transformed dataset

- SAMAX (Score acquired maximum) – 1 if the account achieved a maximum task score, 0 otherwise,
- BA (Buy answer) – 1 if an answer was purchased for the task, 0 otherwise,
- T (Time) - total time spent per task.

Finally, the binary indicator SA (Suspicious Account) flags whether an account meets the thresholds established earlier in Section "Methodology": total score ≤ 100 points and total purchased answers ≤ 10. Thus, SA = 1 for accounts satisfying both criteria, and 0 otherwise.

In contrast to account-level aggregation, the second part of the transformed dataset preserves task-level behavioral measures for each of the 60 suspicious tasks (Fig. 3). To ensure consistency between analytical views, each task is described using the same four characteristics (R, SAMAX, BA, and T) (highlighted in grey background, Fig. 2). To clarify the mapping, characteristic names follow the convention CH_ID_TP, in which:

- CH - measured characteristic (R, SAMAX, BA, and T),
- ID - unique task identifier,
- TP - task type.

Given the distribution of suspicious cases, CAMEO behavior was observed across three assessment formats based on difficulty and design:

- Type 8 (open-ended programming tasks): require full program construction from scratch. Notably, this category contained 52 of the 60 CAMEO cases, reflecting the highest cognitive and interactional demand.
- Type 5 (exact-response code completion): involve short snippet writing, often covered in earlier instruction. Comparatively, this type yielded 4 detected CAMEO cases.
- Type 2 (multiple-choice tasks): the simplest task format. Nevertheless, even within these lower-complexity interactions, 4 CAMEO cases were detected.

Data exploratory

Preliminary statistical inspection shows that the majority of behavioral characteristics used to characterize examined accounts exhibit substantial statistical instability. More precisely, 15 out of 20 characteristics reveal statistically significant outliers (*Grubbs test* > 4.05, $p < 0.05$, Table 1). In parallel with this, pronounced distributional asymmetry is confirmed by large gaps not only between the median and mean, but also between the median and robust approximations of central tendency, including trimmed and Winsorized means.

However, variables representing task success - such as relative score performance (RS_TOTAL, RS_SAMPLE) and maximum-score achievement characteristics (SAMAX_8,

R_2429_8	SAMAX_2429	BA_2429_8	T_2429_8	R_2476_8	SAMAX_2476	BA_2476_8	T_2476_8	R_2525_8	SAMAX_2525	BA_2525_8	T_2525_8	R_2526_8	SAMAX_2526	BA_2526_8	T_2526_8	R_2534_8	SAMAX_2534	BA_2534_8	T_2534_8
10	1	0	466	6	1	0	2063	10	1	0	694	14	1	0	1917	5	1	0	858
20	1	0	1996	11	1	0	1893	11	1	0	1025	7	1	0	609	3	1	0	804
5	1	0	322	3	1	0	1021	6	1	0	1108	5	1	0	1962	2	1	0	837
16	1	0	2139	7	0	0	933	9	0	0	1744	8	1	0	2269	5	1	0	1805
4	1	0	91	15	1	0	835	17	1	0	988	4	1	0	271	31	1	0	1317
2	1	0	31	7	1	0	642					8	1	0	1156	1	1	0	30
4	1	0	481	1	1	0	964	6	1	0	952	2	1	0	618	2	1	0	1716
2	1	0	178	8	1	0	1981									4	0	0	203
21	1	0	2742	9	1	0	2181	21	1	0	2008	11	1	0	1054	3	1	0	1135
4	1	0	677	1	1	0	8									1	1	0	16
7	1	0	470	7	1	0	754					11	0	0	1052	2	1	0	461
7	1	0	299	4	1	0	677	2	1	0	435	3	1	0	574	1	1	0	420

Fig. 3 Characteristics for each task

Table 1 Grubbs outlier test and robust estimates of the mean

Variable	Valid N	Median	Mean	Trimmed mean (5%)	Winsorized mean (5%)	Grubbs Test	p-value
RS_TOTAL	377	0.97	0.89	0.93	0.89	3.67	0.0810
R_TOTAL	377	750.00	769.80	726.70	742.51	4.05	0.0158
BA_TOTAL	377	1.00	2.18	1.66	1.90	9.74	0.0000
TIME_TOTAL	377	43909.00	69102.72	59333.92	64177.46	6.72	0.0000
RS_SAMPLE	377	0.94	0.83	0.86	0.83	2.99	0.9845
R_SAMPLE	377	78.00	116.49	100.31	109.30	4.59	0.0012
BA_SAMPLE	377	0.00	0.75	0.52	0.67	9.12	0.0000
TIME_SAMPLE	377	7636.00	17724.78	14217.52	16298.85	5.95	0.0000
R_8	377	1.06	1.88	1.56	1.73	4.95	0.0002
SAMAX_8	377	0.29	0.34	0.32	0.34	2.19	1.0000
BA_8	377	0.00	0.01	0.00	0.01	10.37	0.0000
T_8	377	128.98	323.83	256.92	296.30	5.99	0.0000
R_5	377	1.75	2.84	2.33	2.56	8.11	0.0000
SAMAX_5	377	0.50	0.51	0.51	0.51	1.15	1.0000
BA_5	377	0.00	0.01	0.00	0.00	8.43	0.0000
T_5	377	48.75	107.67	64.95	76.57	10.75	0.0000
R_2	377	1.00	1.82	1.38	1.58	7.12	0.0000
SAMAX_2	377	0.50	0.53	0.54	0.53	1.16	1.0000
BA_2	377	0.00	0.01	0.00	0.00	9.41	0.0000
T_2	377	37.75	113.20	71.04	87.14	10.70	0.0000

Table 2 Relative measures of variability and moments

Variable	Valid N	Coef.Var.	Coef.Dis.	Skewness	Kurtosis
RS_TOTAL	377	**27.25**	**7.22**	-3.19	8.85
R_TOTAL	377	79.84	120.93	0.88	1.16
BA_TOTAL	377	164.00	300.00	4.05	27.26
TIME_TOTAL	377	115.48	187.65	2.42	9.04
RS_SAMPLE	377	**33.40**	**18.09**	-2.22	3.78
R_SAMPLE	377	114.16	191.03	1.92	4.45
BA_SAMPLE	377	194.16		3.75	22.32
TIME_SAMPLE	377	143.25	284.10	2.33	6.57
R_8	377	127.11	212.67	2.24	6.07
SAMAX_8	377	89.71	187.15	0.58	-0.83
BA_8	377	274.95		5.10	37.28
T_8	377	147.29	310.69	2.35	6.57
R_5	377	132.94	228.57	2.90	14.25
SAMAX_5	377	86.69	200.00	-0.04	-1.80
BA_5	377	514.44		5.62	34.32
T_5	377	255.27	209.74	7.00	59.10
R_2	377	155.64	225.00	3.28	14.20
SAMAX_2	377	85.97	200.00	-0.10	-1.85
BA_2	377	535.49		6.09	40.34
T_2	377	230.49	311.92	6.06	49.99

SAMAX_5, SAMAX_2) - constitute clear exceptions (Table 1). Unlike the remaining indicators, these variables show no evidence of statistically significant extremes (Table 1).

Similarly, from the perspective of variability, nearly all behavioral measures demonstrate high relative dispersion (*Coef.Var.* > 50%, *Coef.Dis.* > 50%), frequently exceeding 100% (Table 2). Again, RS_TOTAL and RS_SAMPLE (shown in bold in Table 2) diverge

from this pattern, presenting relatively controlled variability (< 50%), thereby remaining the only stable account-level measures in the dataset.

The data are too heterogeneous to be meaningfully summarized using traditional central tendency metrics. This is reinforced by concordant normality test results, which reject the null hypothesis (H0: *Data of observed characteristics come from the assumed normal distribution*.) at a 0.001 significance threshold (Kolmogorov-Smirnov $p < 0.01$, Lilliefors $p < 0.01$, and Shapiro-Wilk's W $p < 0.001$, Table 3). Furthermore, the combined evidence from non-zero statistical moments, extreme relative dispersion (Table 2), and frequent outlier presence (Table 1) clearly affirms systematic deviation from normality across observed behavioral characteristics.

These findings lead us to two key observations. First, the empirical quality of MOOC-derived log data is unexpectedly low when evaluated through distributional stability, posing inherent challenges for parametric modeling. Second, and more critically, because normal distribution assumptions are violated across nearly all characteristics, the analysis must shift toward non-parametric statistical methods. Accordingly, with the exception of RS_TOTAL and RS_SAMPLE, we avoid interpretations based on mean values and instead frame account behavior through variability-focused profiling.

Similar to variance-centric anomaly analysis used in industrial quality monitoring, we prioritize robustness to dispersion, behavioral volatility, and non-Gaussian irregularity structures in student interaction signatures. Ultimately, this orientation better reflects the asynchronous and highly uneven interaction dynamics of MOOC platforms and supports the analytical direction adopted in the subsequent stages of our study.

Hypothesis

In light of the extreme heterogeneity and non-normal structure identified during the exploratory phase, we formulate the following testable assumptions. Specifically, our hypotheses center on whether cheating-oriented accounts demonstrate reduced

Table 3 Tests of Normality

	N	max D	K-S p	Lilliefors p	W	p
RS_TOTAL	377	0.337	< 0.01	< 0.01	0.443	0.0000
R_TOTAL	377	0.105	< 0.01	< 0.01	0.932	0.0000
BA_TOTAL	377	0.271	< 0.01	< 0.01	0.612	0.0000
TIME_TOTAL	377	0.193	< 0.01	< 0.01	0.774	0.0000
RS_SAMPLE	377	0.266	< 0.01	< 0.01	0.617	0.0000
R_SAMPLE	377	0.191	< 0.01	< 0.01	0.798	0.0000
BA_SAMPLE	377	0.339	< 0.01	< 0.01	0.562	0.0000
TIME_SAMPLE	377	0.243	< 0.01	< 0.01	0.710	0.0000
R_8	377	0.216	< 0.01	< 0.01	0.750	0.0000
SAMAX_8	377	0.133	< 0.01	< 0.01	0.902	0.0000
BA_8	377	0.422	< 0.01	< 0.01	0.409	0.0000
T_8	377	0.249	< 0.01	< 0.01	0.700	0.0000
R_5	377	0.226	< 0.01	< 0.01	0.724	0.0000
SAMAX_5	377	0.251	< 0.01	< 0.01	0.762	0.0000
BA_5	377	0.537	< 0.01	< 0.01	0.191	0.0000
T_5	377	0.348	< 0.01	< 0.01	0.351	0.0000
R_2	377	0.260	< 0.01	< 0.01	0.640	0.0000
SAMAX_2	377	0.297	< 0.01	< 0.01	0.730	0.0000
BA_2	377	0.534	< 0.01	< 0.01	0.185	0.0000
T_2	377	0.332	< 0.01	< 0.01	0.433	0.0000

behavioral volatility when compared to normally engaged students. Accordingly, we state:

1) We hypothesize that suspicious accounts will show statistically significant lower variability in the behavioral vector representing engagement across the full course (RS_TOTAL, R_TOTAL, BA_TOTAL, and TIME_TOTAL).
2) We hypothesize that suspicious accounts will show statistically significant lower variability in the behavioral vector derived from the 60 selected suspicious tasks where answer purchasing was recorded or expected (RS_SAMPLE, R_SAMPLE, BA_SAMPLE, TIME_SAMPLE).
3) We hypothesize that suspicious accounts will show statistically significant lower variability in the behavioral vector corresponding to task type 8 (SAMAX_8, R_8, BA_8, T_8), which requires development of a full program solution.
4) Furthermore, based on behavioral volatility observed in type 8 tasks, we introduce a set of feature-specific hypotheses and assume that variability-based detection will be effective when applied individually to each type 8 behavioral characteristic. Thus, we posit effectiveness for:

    a. variability of the SAMAX_8 characteristic (maximum-score correctness per type 8 task),
    b. variability of the R_8 characteristic (attempt frequency),
    c. variability of the BA_8 characteristic (use of "Buy Answer"), and
    d. variability of the T_8 characteristic (time on task).

To test these assumptions, and given the distributional violations confirmed earlier, we employ a non-parametric variability analysis strategy rather than mean-based modeling. More precisely, hypothesis verification will rely on:

- Non-parametric univariate and multivariate dispersion tests,
- Robust variability estimators resilient to extreme values, and
- Direct visualization of raw behavioral distributions, to preserve sensitivity to outliers rather than suppress them through parametric assumptions.

These hypotheses reflect a detection philosophy grounded in behavioral consistency rather than response content similarity, aligning our analytical focus with the asynchronous and open-interaction affordances inherent to the platform studied - most notably those observed in the educational MOOC environment of Priscilla.

Results

To begin with the multivariate test results (Table 4) indicate statistically significant differences between suspicious and remaining accounts across the behavioral engagement vector representing all tasks (RS_TOTAL, R_TOTAL, BA_TOTAL, TIME_TOTAL).

Following this, univariate dispersion tests (Table 5) further confirm that these differences persist at the level of individual characteristics.

However, despite the overall significance, the first hypothesis is only partially supported. More specifically, statistically significantly lower variability for suspicious accounts was observed for R_TOTAL, BA_TOTAL, and TIME_TOTAL (Table 5), a trend that is also visually evident in the raw data distributions (Fig. 4b, c, and d).

Table 4 Non-parametric multivariate test of variability and robust estimates of the variability for all task characteristics (total)

	SA = 0			SA = 1		
	Valid *N*	**Aver.Dev.**	**Quar.Dev.**	**Valid *N***	**Aver.Dev.**	**Quar.Dev.**
RS_TOTAL	337	**0.04**	**0.03**	38	0.45	0.50
R_TOTAL	337	457.92	378.00	38	**0.90**	**0.50**
BA_TOTAL	337	2.44	1.50	38	**0.21**	**0.00**
TIME_TOTAL	337	57036.08	40180.00	38	**17.04**	**0.00**

Sen-Puri test: $N = 377$, *Chi-square* = 1568.522, df = 10, $p < 0.05$

Table 5 Non-parametric univariate test of variability for all task characteristics (total)

	Var.Res.	**F**	***p***
RS_TOTAL	0.005	1197.583	0.0000
R_TOTAL	122526.000	58.249	0.0000
BA_TOTAL	7.421	22.931	0.0000
TIME_TOTAL	2906585000.000	38.221	0.0000

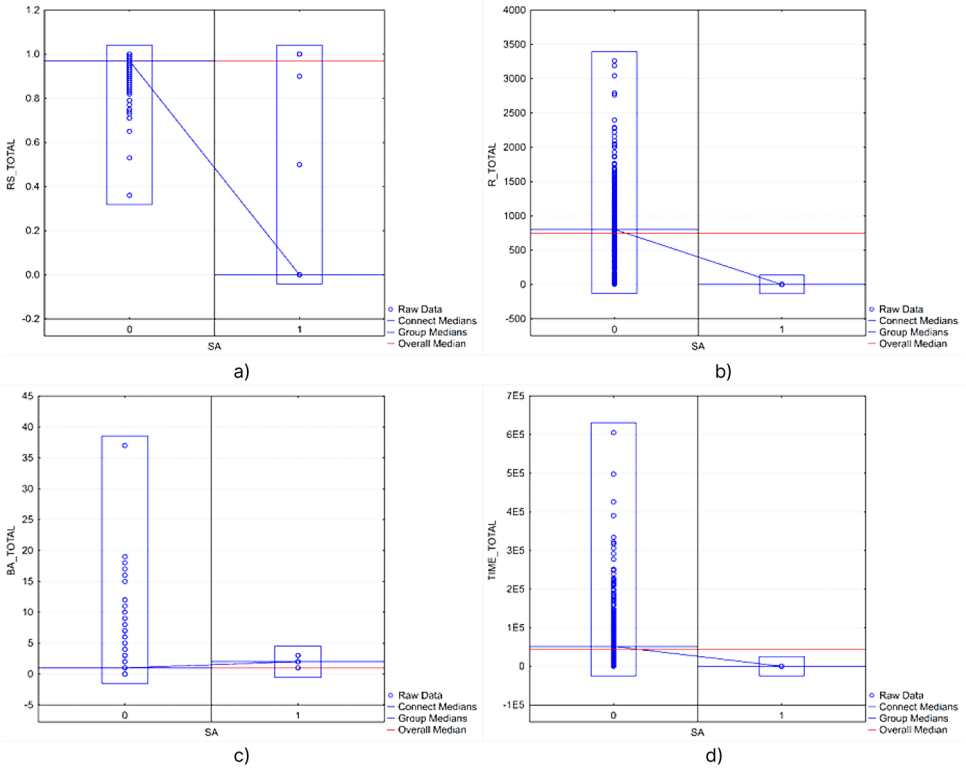

Fig. 4 Raw data plot for all task characteristics: (**a**) RS_TOTAL; (**b**) R_TOTAL; (**c**) BA_TOTAL; (**d**) TIME_TOTAL

By contrast, although RS_TOTAL also demonstrated significant variability differences, lower relative variability was detected in non-cheating accounts instead, indicating an inverse pattern (Table 4; Fig. 4a). This suggests that, unlike other engagement characteristics, RS_TOTAL requires interpretation in conjunction with mean-level performance trends, which echoes observations noted earlier in the exploratory phase.

Similarly, the second multivariate test (Table 6) shows statistically significant differences in variability within the suspicious purchasing-context task vector (RS_SAMPLE, R_SAMPLE, BA_SAMPLE, TIME_SAMPLE).

Table 6 Non-parametric multivariate test of variability and robust estimates of the variability for characteristics of the selected tasks (sample)

	SA = 0			SA = 1		
	Valid *N*	**Aver.Dev.**	**Quar.Dev.**	**Valid *N***	**Aver.Dev.**	**Quar.Dev.**
RS_SAMPLE	337	**0.11**	**0.07**	38	0.43	0.50
R_SAMPLE	337	98.38	72.50	38	**0.76**	**0.50**
BA_SAMPLE	337	0.90	0.50	38	**0.42**	0.50
TIME_SAMPLE	337	18716.78	11763.50	38	**16.91**	**1.00**

Sen-Puri test: $N = 377$, *Chi-square* = 910.710, df = 10, $p < 0.05$

Table 7 Non-parametric univariate test of variability for characteristics of the selected tasks (sample)

	Var.Res.	**F**	***p***
RS_SAMPLE	0.000	190.892	0.0000
R_SAMPLE	7469.000	43.599	0.0000
BA_SAMPLE	1.000	6.498	0.0112
TIME_SAMPLE	294443020.000	40.581	0.0000

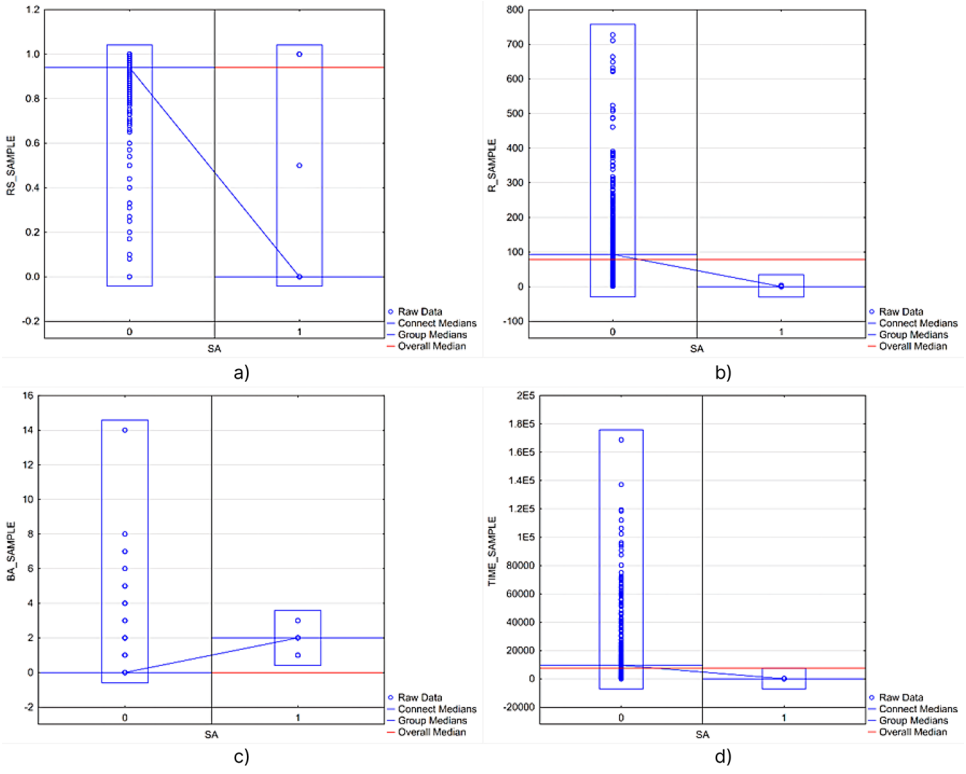

Fig. 5 Raw data plot for characteristics of the selected tasks: (**a**) RS_SAMPLE; (**b**) R_SAMPLE; (**c**) BA_SAMPLE; (**d**) TIME_SAMPLE

Consistent with the previous pattern, univariate dispersion tests (Table 7) also confirm significant differences for each characteristic independently.

Nevertheless, as with Hypothesis 1, Hypothesis 2 is only partially confirmed. Specifically, statistically significantly lower variability for suspicious accounts was observed for R_SAMPLE, BA_SAMPLE, and TIME_SAMPLE, as also reflected in behavioral spread visualizations (Fig. 5b and c, and d). On the contrary, RS_SAMPLE again displayed lower variability only in the remaining (non-cheating) accounts, further supporting the

Table 8 Non-parametric multivariate test of variability and robust estimates of the variability for characteristics of task type 8

	SA = 0			SA = 1		
	Valid *N*	**Aver.Dev.**	**Quar.Dev.**	**Valid *N***	**Aver.Dev.**	**Quar.Dev.**
SAMAX_8	337	0.256	0.250	38	**0.011**	**0.010**
R_8	337	1.731	1.231	38	**0.015**	**0.010**
BA_8	337	0.013	**0.000**	38	**0.012**	0.010
T_8	337	350.91	206.53	38	**0.43**	**0.16**

Sen-Puri test: $N = 377$, *Chi-square* = 1176.098, df = 10, $p < 0.05$

Table 9 Non-parametric univariate test of variability for task characteristics of type 8

	Var.Res.	**F**	***p***
SAMAX_8	0.000	99.235	0.0000
R_8	2.600	38.222	0.0000
BA_8	0.000	0.211	0.6461
T_8	104984.300	39.978	0.0000

Finally, to evaluate individual feature effectiveness, univariate dispersion tests across individual type 8 tasks (Tables 10, 11, 12 and 13) demonstrate the following:

Table 10 Non-parametric univariate test of variability for SAMAX of individual tasks of type 8 (2381, 2396, 2405, 2410, 2429, 2432, 2439, 2441, 2442, 2444, 2446, 2471, 2476, 2507, 2533, 2535, 2554, 2557)

SAMAX	**F**	***p***	**Relative Differentiation**	**Cumulative Differentiation**
2381, 2396, 2405, 2410, 2429, 2432, 2439, 2441, 2442, 2444, 2446, 2471, 2476, 2507, 2533, 2535, 2554, 2557	> 29.592	< 0.001	1.00	1.00

conclusion that relative score metrics exhibit more stability among normally engaged learners (Table 6; Fig. 4a).

these parallel results reveal a meaningful behavioral profile: suspicious accounts demonstrate higher variability in RS (relative score) but lower variability in R, BA, and TIME, regardless of whether all tasks or only suspicious purchasing-context tasks are considered. Thus, variability-based differentiation appears viable when focused on behavioral volatility rather than mean estimates alone (Fig. 5).

Turning to type 8 task behavior, the multivariate variability test (Table 8) also reveals statistically significant dispersion differences between suspicious and remaining accounts along the vector representing full program-writing tasks (RS_8, R_8, BA_8, and TIME_8). In this case, the third hypothesis is again partially valid, and statistically significantly lower variability for suspicious accounts was observed for SAMAX_8, R_8, and T_8.

By contrast, no statistically significant variability difference was found for BA_8 (Table 9), indicating weaker discriminatory power for answer-purchase frequency in this task type at the aggregated level.

- With respect to maximum-score correctness variability (SAMAX; Table 10), accounts were 100% effectively differentiated between suspicious and remaining groups at the 0.001 significance level.

Table 11 Non-parametric univariate test of variability for R of individual tasks of type 8 (2381, 2396, 2405, 2410, 2429, 2432, 2439, 2441, 2442, 2444, 2446, 2471, 2476, 2507, 2533, 2535, 2554, 2557)

R	F	*p*	Relative Differentiation	Cumulative Differentiation
2429, 2476, 2471, 2405, 2444, 2439, 2432, 2446, 2396, 2554, 2557, 2533, 2441, 2507	> 11.302	< 0.001	0.78	0.78
2442, 2535, 2410	> 7.49	< 0.01	0.17	0.94
2381	> 5.117	< 0.05	0.06	1.00

Table 12 Non-parametric univariate test of variability for BA of individual tasks of type 8 (2381, 2396, 2405, 2410, 2429, 2432, 2439, 2441, 2442, 2444, 2446, 2471, 2476, 2507, 2533, 2535, 2554, 2557)

BA	F	*p*	Relative Differentiation	Cumulative Differentiation
2381, 2439, 2507, 2533, 2396, 2405, 2429, 2442	> 14.183	< 0.001	0.44	0.44
2441, 2446, 2535	> 7.107	< 0.01	0.17	0.61
2476	> 3.889	< 0.05	0.06	0.67
2557, 2410, 2554, 2471, 2444, 2432	> 0.168	> 0.05	0.33	1.00

Table 13 Non-parametric univariate test of variability for T of individual tasks of type 8 (2381, 2396, 2405, 2410, 2429, 2432, 2439, 2441, 2442, 2444, 2446, 2471, 2476, 2507, 2533, 2535, 2554, 2557)

T	F	*p*	Relative Differentiation	Cumulative Differentiation
2476, 2439, 2444, 2471	> 12.962	< 0.001	0,22	0,22
2507, 2554, 2432, 2533, 2442, 2441, 2405, 2410, 2557	> 7.971	< 0.01	0,50	0,72
2446, 2396, 2429, 2535, 2381	> 4.466	< 0.05	0,28	1,00

- Similarly, for attempt frequency variability (R; Table 11), effectiveness improves as significance thresholds: 78% at 0.001, 94% at 0.01, and 100% at 0.05.
- For buy answer variability (BA; Table 12), discriminatory effectiveness is lower: 44% at 0.001, 61% at 0.01, and 67% at 0.05. Although weaker than other indicators, only 33% of suspicious accounts remain indistinguishable based on BA variability, which aligns with aggregated findings for BA_8 (Table 9).
- For time on task variability (T; Table 13), effectiveness is limited at strict significance (22% at 0.001) but becomes highly reliable at 0.01 (72%) and fully effective at 0.05 (100%).

Variability of the maximum-score correctness (SAMAX_8), attempt frequency (R_8), and time engagement (T_8) characteristics reached 100% discrimination effectiveness at the 0.05 level, while BA_8 reached only 67%, marking it as the weakest characteristic for type 8.

Our hypotheses are consistently partially validated, and the findings collectively demonstrate that fraudulent multi-account cheating accounts exhibit distinctive variability compression patterns, particularly in maximum-score correctness, attempt frequency, and time engagement behavior. These results confirm the suitability of non-parametric, variance-centric modeling for anomaly-driven academic misconduct detection within the examined MOOC environment of Priscilla.

Discussion

To begin with, and in alignment with our stated goals, the primary objective of this study was to propose a new method for detecting cheating behavior in MOOCs, specifically targeting CAMEO-style cheating through the identification of suspicious harvester and master accounts. Although our approach was inspired by Northcutt et al. (2016), it departs in a key respect by deliberately excluding IP-address matching, given that such linkage has been shown to be unreliable in shared or masked network environments, including institutional Wi-Fi, VPN routing, and proxy-based connections.

More specifically, in contrast to IP-driven account filtering, our filter-based algorithm instead relies on temporal precedence and behavioral dispersion patterns, especially those associated with Immediate Mode harvesting. By adopting this adjustment, we were able to detect a greater number of suspicious harvester accounts without generating false positives induced by shared institutional network infrastructure. In doing so, we identified that approximately 6.8% of course accounts engaged with the platform-specific "Buy Answer" feature, purchasing correct responses that were subsequently used for submission in a linked master account. That said, and as expected, this proportion may be understated because batch-mode CAMEO behavior was not fully traceable under the current linkage window, meaning that accounts harvesting answers for delayed reuse were not always linkable by timing alone.

For contextual comparison, prior studies confirm that MOOC cheating via multi-account harvesting is neither rare nor uniform across courses. For example, Northcutt et al. (2016) and Bao et al. (2017) applied filter-based detection to multiple MOOCs and found that between 1.3% and 1.89% of certificates were earned through CAMEO interactions at the aggregate level, while specific single-course prevalence could reach up to 7% (Bao et al. 2017). Likewise, complementary works by Ruiperez-Valiente et al. (2016) and Alexandron et al. (2017) show that between 10% and 12.9% of certificate holders used CAMEO to source at least 1% of their correct responses through harvested answers.

Nonetheless, because the Priscilla MOOC platform does not issue certificates, the educational impact of cheating must be interpreted at the institutional course-completion level rather than certificate fraud. From this perspective, therefore, we estimate that approximately 1.1% of course-completing students engaged in CAMEO-style cheating. Notably, and consistent with prior MOOC observations, harvesting activity was detected across 60 unique assessment tasks, representing approximately 15% of all course tasks, of which 86% belonged to task type 8 - the most demanding format requiring full program construction. This aligns with the findings of Bao et al. (2017), who emphasized that CAMEO behavior disproportionately targets higher-difficulty tasks. In our data, specifically, the dominance of type 8 tasks (52 of 60 cases) further supports the conclusion that cheating likelihood increases alongside task complexity.

Consequently, following the identification stage, our analysis shifted toward evaluating behavioral variability as a detection signal, particularly because traditional measures of central tendency were untenable given the heavy-tailed distributional structure confirmed in Sect. "Data exploratory" This led us to observe that suspicious harvester accounts differ sharply from normally engaged students not only in restricted engagement profiles (≤ 100 total points and ≤ 10 answer purchases) but also in behavioral consistency compression, exhibiting: (a) very few repeated task attempts, (b) short total time spent interacting with assessments, and (3) frequent reliance on the "Buy Answer"

feature for maximum-score correctness extraction. By contrast, while their engagement variability is compressed, relative score variability (RS_TOTAL and RS_SAMPLE) remains high, spanning 0.0–1.0, whereas normally engaged accounts show tighter clustering (0.36-1.0).

Furthermore, to validate feature effectiveness, task-level variability tests (Tables 10, 11, 12 and 13) confirmed that the most differentiated characteristics include: maximum acquired score, repetitions (number of task repetitions), and time (time spent on task) behaviors, a finding that corresponds closely with prior MOOC profiling outcomes (Alexandron et al. 2016, 2019; Ruiperez-Valiente et al. 2017). These accounts are characterized by spending short time in the course, while their accuracy—measured as the relative score in our case—is either very high (1.0) or very low (0.0). For tasks where the answer was purchased, the harvester account could either submit the previous bought answer or leave the question unanswered until they will obtain the answer.

Moreover, although temporal linkage handled immediate behavior effectively, batch mode detection remains limited to partial pattern recognition, not high confidence linkage. Although behavioral characteristics remain similar, delayed submission and many to one harvester master dependency obscure deterministic pairing, as multiple master accounts may benefit from the same harvester. Thus, accurate batch mode account linking remains an open methodological challenge.

Finally, we agree with Monahan and Shah (2023) that academic dishonesty extends beyond the classroom. Although we cannot eliminate unethical intent, institutions can minimize opportunity, enhance accountability, and reinforce deterrence, particularly through deliberate platform security design, integrity policy enforcement, and AI-ethical usage norms. Ultimately, therefore, detection models grounded in behavioral variability - rather than network metadata - offer a promising and scalable pathway for MOOC cheating detection in asynchronous educational environments, including on platforms such as the educational micro-learning MOOC provided by Priscilla.

Conclusion

In response to the challenges outlined at the beginning of this paper, the rise of online education has amplified concerns surrounding academic integrity, especially as opportunistic cheating strategies such as CAMEO become increasingly sophisticated in MOOC environments. As previous research has established, trust remains foundational for sustaining educational quality, protecting institutional investments, and ensuring long-term adoption of digital learning models (Anand Shankar Raja and Kallarakal 2021; Bernal Arellano et al. 2024; Chen et al. 2020; Maatuk et al. 2022; Turnbull et al. 2021). Furthermore, the nature of online assessment highlights how traditional cheating behaviors may creatively become embedded within learning processes, requiring a re-examination of educational and assessment norms (Vaezi et al. 2024).

This study has demonstrated that cheating-detection systems can operate effectively without reliance on IP address tracking, a dependency that, although once common, is now vulnerable to misclassification in shared or masked networks. By contrast, and as our results confirm, placing analytic emphasis on temporal precedence and behavioral variability patterns enables more reliable identification of suspicious harvester accounts. Specifically, harvester accounts were distinguished by low variability in engagement characteristics - *task repetitions*, *answer purchases*, and *time spent on tasks* and a higher

variability in *relative scores* compared to other accounts. These opposing stability and volatility dynamics indicate intentional misuse rather than learning inconsistency, a conclusion that aligns with the argument that process-level logs capture behavioral intent more precisely than latent-characteristic or item-misfit models alone (Ranger et al. 2020).

this work bridges a methodological gap by combining filter-based algorithms with temporal sequencing and behavioral analytics, yielding a detection framework capable of uncovering deceptive response-sourcing patterns in asynchronous micro-learning platforms. Importantly, this approach enhances detection sensitivity while also preserving user privacy and reducing metadata-induced noise, thus offering an ethical and scalable alternative for institutions deploying large-scale online learning.

Future work should focus on identifying submission timing clusters, where users submit answers in close temporal proximity, potentially indicating coordinated cheating. Investigating these clusters may reveal collaborative cheating strategies, including immediate mode CAMEO-style activities. To improve batch mode detection, further analysis of submission timestamps, particularly cases involving rapid submissions with high scores, will be essential.

Author contributions
MV and DM conceived this research and designed experiments; MV and DM participated in the design and interpretation of the data; MM performed statistical experiments and data analysis; MV and DM wrote the paper and all authors participated in the revisions of it. All authors read and approved the final manuscript.

Funding
This This work was supported by the Nitra Cybersecurity Competence Center (NKCKB) under project No. 17R05-04-V01-00003, funded by the Recovery and Resilience Plan of the Slovak Republic, within the call 17R05-04-V01 – Establishment of Cybersecurity Competence Centers.

Data availability
The datasets used and/or analysed during the current study are available from the corresponding author on reasonable request.

Declarations

Ethics approval and consent to participate
The study was conducted in accordance with the ethical standards of the institutional research committee and with the 1964 Declaration of Helsinki and its later amendments or comparable ethical standards. All participants provided informed consent at the time of registration to the virtual environment (VE) platform, agreeing to the use and publication of anonymized data and research findings derived from their participation in the study.

Competing interests
The authors declare no competing interests.

Received: 18 June 2025 / Accepted: 2 January 2026
Published online: 22 January 2026

References

Alexandron G, Ruipérez-Valiente JA, Pritchard DE et al (2015). Evidence of MOOC students using multiple accounts to harvest correct answers. Learning with MOOCs II: A Workshop for Practitioners: New Approaches to Teaching & Learning
Alexandron G, Lee S, Chen Z, Pritchard DE (2016) Detecting cheaters in MOOCs using item response theory and learning analytics. CEUR Workshop Proceedings 1618
Alexandron G, Ruipérez-Valiente JA, Chen Z, Muñoz-Merino PJ, Pritchard DE (2017) Copying@Scale: using harvesting accounts for collecting correct answers in a MOOC. Computers Educ 108. https://doi.org/10.1016/j.compedu.2017.01.015
Alexandron G, Ruipérez-Valiente JA, Pritchard DE (2019) Towards a general purpose anomaly detection method to identify cheaters in massive open online courses. EDM 2019 - Proceedings of the 12th International Conference on Educational Data Mining
Anand Shankar Raja M, Kallarakal TK (2021) COVID-19 and students perception about MOOCs a case of Indian higher educational institutions. Interact Technol Smart Educ 18(3). https://doi.org/10.1108/ITSE-07-2020-0106
Aruğaslan E (2024) Examining the relationship of academic dishonesty with academic procrastination, and time management in distance education. Heliyon 10(19):e38827. https://doi.org/10.1016/J.HELIYON.2024.E38827

Balderas A, Caballero-Hernández JA (2020) Analysis of learning records to detect student cheating on online exams: case study during COVID-19 pandemic. ACM Int Conf Proceeding Ser. https://doi.org/10.1145/3434780.3436662
Balderas A, Palomo-Duarte M, Caballero-Hernández JA, Rodriguez-Garcia M, Dodero JM (2021) Learning analytics to detect evidence of fraudulent behaviour in online examinations. Int J Interact Multimedia Artif Intell 7(2). https://doi.org/10.9781/ijimai.2021.10.007
Bao Y, Chen G, Hauff C (2017) On the prevalence of multiple-account cheating in massive open online learning: A replication study. Int Edu Data Min Soc. https://files.eric.ed.gov/fulltext/ED596613.pdf
Bernal Arellano WM, Parra G, Reinoso JNV, J. M., Quito Ochoa JF (2024) A study on academic dishonesty among english as a foreign Language students. Heliyon 10(13):e33876. https://doi.org/10.1016/J.HELIYON.2024.E33876
Caspari-Sadeghi S (2022) Applying learning analytics in online environments: measuring learners' engagement unobtrusively. Front Educ 7. https://doi.org/10.3389/feduc.2022.840947
Chen C, Long J, Liu J, Wang Z, Wang L, Zhang J (2020) Online academic dishonesty of college students: a review. https://doi.org/10.2991/assehr.k.200723.121
Colnerud G, Rosander M (2009) Academic dishonesty, ethical norms and learning. Assess Evaluation High Educ 34(5). https://doi.org/10.1080/02602930802155263
Crittenden VL, Hanna RC, Peterson RA (2009) The cheating culture: A global societal phenomenon. Bus Horiz 52(4). https://doi.org/10.1016/j.bushor.2009.02.004
Du J, Song Y, An M, An M, Bogart C, Sakr M (2022) Cheating detection in online assessments via timeline analysis. SIGCSE 2022 - Proc 53rd ACM Tech Symp Comput Sci Educ 1. https://doi.org/10.1145/3478431.3499368
Eckstein M (2003) Combating academic fraud: towards a culture of integrity. https://unesdoc.unesco.org/ark:/48223/pf0000133038
Erguvan ID (2022) University students' Understanding of contract cheating: a qualitative case study in Kuwait. Lang Test Asia 12(1). https://doi.org/10.1186/s40468-022-00208-y
Hasri A, Supar R, Azman NDN, Sharip H, Yamin LSM (2022) Students' attitudes and behavior towards academic dishonesty during online learning. https://doi.org/10.3390/proceedings2022082036
Jaramillo-Morillo D, Ruipérez-Valiente J, Sarasty MF, Ramírez-Gonzalez G (2020) Identifying and characterizing students suspected of academic dishonesty in SPOCs for credit through learning analytics. Int J Educational Technol High Educ 17(1). https://doi.org/10.1186/s41239-020-00221-2
Jenkins BD, Golding JM, Grand L, Levi AM, M. M., Pals AM (2023) When opportunity knocks: college students' cheating amid the COVID-19 pandemic. Teach Psychol 50(4). https://doi.org/10.1177/00986283211059067
Lingán-Huamán SK, Dominguez-Lara S, Carranza Esteban RF (2024) Gender-based differences in the impact of dark triad traits on academic dishonesty: the mediating role of moral disengagement in college students. Heliyon 10(1). https://doi.org/10.1016/j.heliyon.2023.e23322
Maatuk AM, Elberkawi EK, Aljawarneh S, Rashaideh H, Alharbi H (2022) The COVID-19 pandemic and E-learning: challenges and opportunities from the perspective of students and instructors. J Comput High Educ 34(1). https://doi.org/10.1007/s12528-021-09274-2
Monahan M, Shah A (2023) Academic dishonesty: an exploratory study of traditional versus non-traditional students. Res High Educ J 43
Noorbehbahani F, Mohammadi A, Aminazadeh M (2022) A systematic review of research on cheating in online exams from 2010 to 2021. Educ Inform Technol 27(6). https://doi.org/10.1007/s10639-022-10927-7
Northcutt CG, Ho AD, Chuang IL (2016) Detecting and preventing multiple-account cheating in massive open online courses. Computers Educ. 100https://doi.org/10.1016/j.compedu.2016.04.008
Odongo DA, Agyemang E, Forkuor JB (2021) Innovative approaches to cheating: an exploration of examination cheating techniques among tertiary students. Education Research International. https://doi.org/10.1155/2021/6639429
Ranger J, Schmidt N, Wolgast A (2020) The detection of cheating on E-Exams in higher education—The performance of several old and some new indicators. Front Psychol 11. https://doi.org/10.3389/fpsyg.2020.568825
Ruiperez-Valiente JA, Alexandron G, Chen Z, Pritchard DE (2016) Using multiple accounts for harvesting solutions in MOOCs. L@S 2016 - Proc 3rd 2016 ACM Conf Learn Scale. https://doi.org/10.1145/2876034.2876037
Ruiperez-Valiente JA, Munoz-Merino PJ, Alexandron G, Pritchard DE (2017) Using machine learning to detect 'Multiple-Account' cheating and analyze the influence of student and problem features. IEEE Trans Learn Technol 12(1):112–122. https://doi.org/10.1109/TLT.2017.2784420
Ruipérez-Valiente JA, Gašević D, Joksimović S, Kovanović V, Muñoz-Merino PJ, Kloos CD (2017) A data-driven method for the detection of close submitters in online learning environments. 26th Int World Wide Web Conf 2017 WWW 2017 Companion. https://doi.org/10.1145/3041021.3054161
Salamah I (2022) Factors influencing academic dishonesty among Sriwijaya state Polytechnic students. Int J Sci Technol Manage 3(2). https://doi.org/10.46729/ijstm.v3i2.491
Salehi M, Gholampour S (2021) Cheating on exams: investigating reasons, attitudes, and the role of demographic variables. SAGE Open 11(2). https://doi.org/10.1177/21582440211004156
Salih S, Husain O, Hamdan M, Abdelsalam S, Elshafie H, Motwakel A (2025) Transforming education with AI: a systematic review of chatgpt's role in learning, academic practices, and institutional adoption. Results Eng 25:103837. https://doi.org/10.1016/J.RINENG.2024.103837
Skalka J, Drlík M (2018) Priscilla – proposal of system architecture for programming learning and teaching environment. 2018 IEEE 12th International Conference on Application of Information and Communication Technologies (AICT) 1–6. https://doi.org/10.1109/ICAICT.2018.8746921
Sorour SE, AlBarrak KM, Abohany AA, El-Mageed AAA (2024) Credit card fraud detection using the brown bear optimization algorithm. Alexandria Eng J 104:171–192. https://doi.org/10.1016/J.AEJ.2024.06.040
Sozon M, Mohammad Alkharabsheh OH, Fong PW, Chuan SB (2024) Cheating and plagiarism in higher education institutions (HEIs): A literature review. F1000Research 13:788. https://doi.org/10.12688/f1000research.147140.2
Taşkın N (2024) Cheating and prevention strategies in online assessment. 157–178. https://doi.org/10.4018/979-8-3693-3045-6.ch009
Taşkın N, Kokoç M (2025) Behavioural engagement, academic dishonesty, and performance gaps: comparing online and paper–pencil based tests in an online learning context. Educ Inform Technol. https://doi.org/10.1007/s10639-025-13514-8

Tertiary Education Quality and Standards Agency (2025) Student academic misconduct resources – glossary. https://www.teqsa.gov.au/students/student-academic-misconduct-resources/glossary
Turnbull D, Chugh R, Luck J (2021) Transitioning to E-Learning during the COVID-19 pandemic: how have higher education institutions responded to the challenge? Educ Inform Technol 26(5). https://doi.org/10.1007/s10639-021-10633-w
Vaezi S, Vaezi M, Nami F (2024) Academic dishonesty out, use of resources in. Computers Educ Open 6:100193. https://doi.org/10.1016/J.CAEO.2024.100193
Wang H, Dang A, Wu Z, Mac S (2024) Generative AI in higher education: seeing ChatGPT through universities' policies, resources, and guidelines. Computers Education: Artif Intell 7:100326. https://doi.org/10.1016/J.CAEAI.2024.100326
Wu J, Hu R, Li D, Ren L, Hu W, Zang Y (2024) A GNN-based fraud detector with dual resistance to graph disassortativity and imbalance. Inf Sci 669:120580. https://doi.org/10.1016/J.INS.2024.120580
Yu S, Assor A, Yu H, Wang Q (2023) Autonomy as a potential moral resource: parenting practices supporting youth need for autonomy negatively predict youth acceptance of academic dishonesty. Learn Individual Differences 101. https://doi.org/10.1016/j.lindif.2022.102245
Yuan J, Qiu X, Wu J, Guo J, Li W, Wang Y-G (2024) Integrating behavior analysis with machine learning to predict online learning performance: a scientometric review and empirical study. http://arxiv.org/abs/2406.11847
Zhao L, Mao H, Compton BJ, Peng J, Fu G, Fang F, Heyman GD, Lee K (2022) Academic dishonesty and its relations to peer cheating and culture: A meta-analysis of the perceived peer cheating effect. Educational Res Rev 36. https://doi.org/10.1016/j.edurev.2022.100455

Publisher's note