

# Remote Proctored Exam for Proficiency-Testing: A Cross-Sectional Study

Vasiliki Andreou (✉ [vasiliki.andreou@kuleuven.be](mailto:vasiliki.andreou@kuleuven.be))

KU Leuven

Sanne Peters

KU Leuven

Jan Eggermont

KU Leuven

Johan Wens

University of Antwerp

Birgitte Schoenmakers

KU Leuven

---

## Research Article

**Keywords:** General Practice, Postgraduate Education, Summative evaluation, Online assessment

**Posted Date:** March 18th, 2021

**DOI:** <https://doi.org/10.21203/rs.3.rs-296304/v1>

**License:**   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

---

# Abstract

**Background:** The COVID-19 pandemic has profoundly affected assessment practice in higher education including a complex planning of supervision. To organise safely and reliably a remote proficiency-test for admission to the Advanced Master of General Practice (AMGP), we developed a supervisor app tracking and tracing candidates' behaviour.

**Methods:** A cross-sectional design was adopted with candidates applying for admission to the AMGP. The supervisor app operated on three levels to register events: recording actions, analysing behaviour, and live supervision. Each suspicious event was given a score. The outcome measures were the number of suspicious events and the exam outcome compared to the past year. To get more insight into candidates' perceptions about the app, a post-test questionnaire was administered. An exploratory factor analysis was performed to explore quantitative data, while qualitative data were thematically analysed.

**Results:** In total, 472 (79%) candidates used the app in an off campus setting and 121 (20%) were on campus with live supervision. Test results of both groups were comparable. The app detected 22 candidates with a suspicious level  $>1$ , mainly due to background noise. All events occurred without fraud purpose. Out of 472 candidates, 304 filled in the post-test questionnaire. Two factors were extracted from the analysis and identified as candidates' appreciation of the app and as emotional distress because of the app. Four themes were identified in the thematic analysis providing more insight on candidates' emotional well-being.

**Conclusions:** A supervisor app registering and recording behaviour to prevent fraud during off-campus exams is efficient without influencing the exam outcome. Although candidates' perceptions were mixed, increased anxiety was due to the lack of clear guidelines about the app. Future research should compare in a controlled design the cost-benefit balance between the supervisor app and candidates' awareness of being monitored combined with a safe exam browsing plug in.

## Introduction

The COVID-19 pandemic has profoundly affected teaching and assessment programs at all levels of education. The Flemish (Belgium) universities were challenged to switch fluently from analogous to digital mode which went with ups and downs. Moreover, the COVID-19 countermeasures require sufficient physical distance before and during exams necessitating smaller student groups per exam session. This situation demanded complex planning of manpower, location, and, in the case of digital exams, adequate equipment, both hard and software.

The challenge of reorganization is even more impressive in an interuniversity collaboration for education. The Advanced Master's in General Practice (AMGP) in Flanders, Belgium, is formally organized and offered by four Flemish universities (KU Leuven, University of Antwerp, University of Ghent, VUB). This collaboration comprises common administration, curriculum, examinations, and residencies, but separate registration of residents to the university of their choice. Given, we have over 900 residents in the AMGP

training and education. Exams planning is a complex logistical and administrative process. Therefore, we built more than a decade ago an intelligent, comprehensive, and interactive assessment platform. Medical candidates are in favour of computerized exams and these exams may enhance the learning experience and effect.(1) Our platform offers the interface for summative and formative knowledge testing (in six questions formats), Objective Structured Clinical Examination (OSCE)-performance, and proficiency testing.

Proficiency-testing takes place outside the regular exam regulations and is organised for the four universities together. The test comprises three stages starting with an administrative stage, followed by an actual exam, and finalized by a jury exam for candidates who failed the exam in stage two. The actual exam is an online machine-assisted test that runs on the digital assessment platform and consists of three components: knowledge testing, critical reasoning testing and situational judgment testing. The whole procedure runs since 2016 and has proved its reliability, validity, acceptability, and feasibility in this format.(2, 3)

The necessity of respecting COVID-19 countermeasures but adhering to the original exam format forced us to adopt a creative solution. Therefore, we decided and organised a remote proctored exam in addition to the on-campus exam. To prevent fraud, we developed a virtual supervisor app tracking and tracing candidates' behaviour during the exam. The technology we built and applied goes beyond the traditional proctored exam systems where focus lies on recording sound and image.(4) In this article, we report the outcome of the anti-fraud measures and compare the scores between off- and on-campus exams. This study also aims at providing insights about candidates' perceptions about the supervisor app.

## Methods

In collaboration with the developers of the assessment platform and in discussion with the coordinators of the AMGP, we determined the criteria and conditions to design the supervisor app based on our online assessment platform. That implies that the app is not immediately available for third parties because of potential compatibility issues. Specifically, we integrated an Application Programming Interface software (Vonage APIs) within our assessment platform. The Vonage APIs software enabled to and interact with and record the candidates live. Along with this software, we implemented several metrics to detect events implicating suspicious behaviour (Figure 1). Suspicious events were defined as: switch to another browser, return to page, close page, disconnection from internet and sound/noise.

During the exam, the system recorded three channels: the computer screen, the camera, and the microphone. The supervisor app operated on three levels: recording of actions, analysis of behaviour and live supervision. These recordings were immediately encrypted and saved on a secured server. Additionally, the app used an algorithm of pattern recognition in responses, clicking behaviour and timestamp analysis. It also analysed both individual behaviour as well as correlations across candidates. Each event was given a score. Every suspicious event increased the result by 0.1. A candidate scoring 0.5 or higher was considered suspicious. The tracking was stored in the users' browser and sent to the server

after each 20 non-suspicious events or when suspicious event occurred. When the suspicious rating was above 0.5, the exam submission was flagged and a report on the suspicious behaviour was downloaded for assessment.

Remote supervision allowed for optional human oversight during the examination. The human monitor could immediately join the live feed of each candidate to obtain more information or to send a warning in a private message. In case of a crash of the supervisor app, the affected candidates were switched to Safe Exam Browser (SEB), which was integrated in our assessment platform. SEB allowed only the exam interface to remain accessible on the machine.

In addition to the technical solution, we expected the app to have an impact on candidates' behaviour regarding fraud prevention. Candidates were comprehensively briefed in advance on how to install the app, on how to test it, and on the specific features of the supervising technology by an animation movie.

A voluntary panel tested the supervisor app in two sessions. During the first session, the participants were ordered to behave in a suspicious manner: talking, making noise, turning away from the screen, using the internet, etc. Afterwards, we made the following adjustments to the app and the procedure: we isolated the sound-suspicious level from the other suspicious events, because of sensitivity of sound detection. We also increased the overall suspicious level from 0.5 to 1 to take the sound sensitivity into account.

The intervention took place during the regular exam period of the proficiency test for admission to the AMGP. All participants were candidates applying for admission. To take the exam, candidates were free to register for either an off- or on-campus participation. On campus, a human supervisor was present, and candidates used the campus gear. Candidates that chose an off-campus option could simultaneously take the exam. For remote monitoring, we engaged an experienced supervising team of six staff members. The off-campus supervisors were able to send a notification or warning to candidates who were behaving suspiciously, and they intervened in case of technical issues. The developers of the app were also fully available during the exam to provide technical assistance, if necessary.

The actual exam and the associated procedures were set up as in previous years: all candidates completed the same exams and candidates who failed the machine-assisted test (the actual exam) were invited to a jury exam one week later. Candidates for whom suspicious behaviour was flagged during the exam or in post-exam analyses were also invited to the jury exam.

To get an insight on candidates' perceptions about the supervisor app, we administered a post-test survey in the form of an online questionnaire. The questionnaire was sent only to the candidates participating in the off-campus exam. Filling the questionnaire was on anonymous and voluntary basis. Candidates were asked whether they were taking the proficiency test for the first time to avoid any effects due to retakes (e.g. test anxiety). Respondents had to specify their level of agreement or disagreement in a 6-point Likert scale for six items regarding the supervisor app. In addition, candidates could explain how they perceived the influence of the app on their exam experience and exam outcome. Table 1 displays a list of the survey questions. All methods were carried out in accordance with relevant guidelines and

regulations. Ethical approval was granted by the Social and Societal Ethics Committee of the KU Leuven with the following approval number: *G-2020-2262-R2(MAR)*.

**Table 1** Structure of the questionnaire measuring candidates' perceptions about supervisor app

<b>Closed questions (6-point Likert scale)</b>
Q1: I was well informed about the app
Q2: I was more nervous than usual before the exam because of the supervisor app
Q3: I was more nervous than usual during the exam because of the supervisor app
Q4: The supervisor app had an impact on my results.
Q5: I found the supervisor app reassuring.
Q6: I would use the supervisor app again in the future for other exams.
<b>Open questions</b>
Q7: How did the supervisor app influence you when taking the exam?
Q8: How do you think that the supervisor app influenced your result?

## Analysis

Outcome measures were defined as the number of suspicious events per student and exam outcome compared to the results of past year. To analyse the post-test questionnaire, we performed an exploratory factor analysis to understand patterns related to candidates' perceptions while using the supervisor app. (5) We used an oblique rotation, since we expected some correlation among factors. (6) The reliability of the scale was calculated based on Cronbach's alpha. (7) We analysed quantitative data with SPSS 27 (IBM SPSS Statistics 27). Qualitative data were thematically analysed by two researchers separately (VA and BS). (8) Discrepancies in coding were discussed until consensus was reached. Data from open-ended questions were analysed in the software program QSR International's NVIVO (Release 1.0).

## Results

A total of 593 candidates subscribed to the exam in the four Flemish universities. Four hundred seventy-two (79%) candidates used the supervisor app for an off-campus exam and 121 (20%) were present on the campus (Table 2). Most candidates registered at the KU Leuven (227) and at the University of Ghent (203) chose to take the exam off-campus. The results of both (off- and on-campus) groups were comparable.

Table 2 Participation and test result off versus on campus

	Off campus n (%)	On campus n (%)	TT-test pooled
Total number of candidates (n=593)	472 (79,6%)	121 (20,4%)	
Number of candidates per university			
- Leuven			
- Antwerp			
- Brussels	227 (84,1%)	43 (15,9%)	
- Gent	29 (35,8%)	52 (64,2%)	
	13 (50%)	13 (50%)	
	203 (94%)	13 (6%°)	
Average test result	72/100	72,8/100	P > 0,15

Overall, we registered and solved 15 technical issues in the off-campus context (Table 3). Eight of these issues concerned software problems (in particular loading a reading text in a new tab). Two candidates experienced a negative impact on the exam performance due to technical issues. The developers team switched one candidate to the SEB mode to complete the exam. Based upon the post exam analyses and after deliberation, the exam coordinator exempted the other candidate of the jury exam.

Table 3 Comparison of exam procedure and outcome off- versus on-campus

	Off campus	On Campus
<b>Technical issues</b>	15	1
- with impact on exam	2	0
<b>Type of issue</b>		
- Internet failure	2	0
- Hardware issue	4	1
- Camera crash	1	0
- Software issue	8	0
<b>Average suspicious score</b>	0,4	NA
<b>Median suspicious score</b>	0,3	NA
<b>Number of suspicious candidates</b>		
- Detected by the app	22 (4%)	NA
- Flagged by supervisors	2 (0,04%)	NA
- With non-critical events <1	455 (96%)	NA
- Without events	15 (3%)	NA
- With noise event	472 (100%)	NA
<b>Interventions during exam</b>		
- Technical intervention	8	1
- Warning to candidate	2 individuals	0
	1 group (background noise)	0

In total, the app detected 22 (4%) candidates with a suspicious level >1. All cases concerned one or more noise-event (background noise). All other non-critical events consisted of leaving the webpage, closing a page or typing text. Live monitoring and a post exam review of records revealed that all these events occurred hazardingly but without fraud purpose.

The monitoring supervisors flagged two candidates who were typing more than expected (in a multiple-choice exam). After revision of the records, these candidates were using 'control find' to search for words

in the reading text. During the exam, the supervisors intervened eight times for a technical issue, they warned two candidates to stop talking to themselves and they sent a group message to ask to turn down background noise.

Out of the 472 candidates that used the supervisor app, 304 filled in the post-test questionnaire, 213 women and 91 men. All of them were taking the proficiency test for the first time. Almost seventy-two percent of the candidates (219) had chosen the GP Training as first choice postgraduate medical training (Figure 2). Ten percent (30) wanted to follow a medical training in Internal Medicine, but they were participating in the proficiency test to be able to register in the AMGP as a second choice.

An explanatory factor analysis was initially conducted on the 6 items of the questionnaire with oblique rotation. However, one item “I was well informed about the supervisor app” had to be omitted because of a communality lower than .40.(6) The other five items (communality > 0.40) were included in a secondary analysis. This analysis yielded a Kaiser-Meyer-Olkin (KMO) of .63 verifying the sampling adequacy (Table 4). After analysing the data to obtain eigenvalues for each factor, two factors had eigenvalues over Kaiser’s criterion of 1 explaining 82.46% of the variance. The scree plot also justified retaining two factors (Figure 3). Table 5 shows the factor loadings after rotation. The items that cluster on the same factors suggest that factor 1 represents candidates’ appreciation of the supervisor app, while factor 2 represents emotional distress because of the supervisor app. Our sample size was appropriate to perform a factor analysis, since it suffices the 10:1 rule of thumb subject to item ration.(6) The reliability of the questionnaire was calculated based on Cronbach’s alpha as 0.72 (Table 6).

Table 4 KMO and Barlett's test

<b>KMO and Barlett’s Test</b>	
Kaiser-Meyer-Olkin measure of sampling adequacy	0.632
Barlett’s Test of sphericity	
Approx. chi-square	667,418
df	10
Sig.	.000

Table 5 Summary of exploratory factor analysis results for the supervisor app questionnaire (n=304)

<b>Rotated Factor Loadings</b>		
Item	Appreciation of the supervisor app	Emotional distress because of the supervisor app
<i>Q2: I was more nervous than usual before the exam because of the supervisor app.</i>	-0.017	<b>.851</b>
<i>Q3: I was more nervous than usual during the exam because of the supervisor app.</i>	-.116	<b>.963</b>
<i>Q4: The supervisor app had an impact on my results.</i>	.121	<b>.692</b>
<i>Q5: I found the supervisor app reassuring.</i>	<b>.973</b>	.035
<i>Q6: I would use the supervisor app again in the future for other exams</i>	<b>.738</b>	-.023
<i>Eigenvalue</i>	2.450	1.673
<i>% of Total Variance</i>	49%	33.46%

Note: Factor loadings over 0.40 appear in bold.

Table 6 Reliability coefficient of the questionnaire

<b>Reliability coefficient</b>	
Cronbach's alpha	.723

From the qualitative analysis, four different themes were discerned. All themes were related to candidates' emotional well-being. The first two themes referred to stress and anxiety before and during the exam. Some candidates felt more anxious because they feared technical problems or running out of time. The third theme was anxiety and stress because of the supervisor app. Candidates mentioned that they felt stress because the app might detect something as fraud, without the intention of fraud. Candidates also admitted that they felt awkward, observed, and distracted because of the app.

Instead of just thinking about a question and the possible answers, I had to constantly remind myself about my behaviour (not to look upwards or left and right). The feeling of constantly being watched was not helpful (Anonymous candidate).

The fourth theme was connected to positive emotions about the supervisor app. Candidates thought that the supervisor app was reassuring and that, in case of need, someone could intervene and help.

The app worked pretty well for me, it was rather reassuring that it would be taken into account, if technical problems arose (Anonymous candidate).

Finally, candidates identified some technical issues before and during taking the exam. Before the exam, the main problem was slow internet connection. During the exam, candidates experienced problems, when opening multiple tabs at the same time.

## Discussion

This pilot study demonstrated that a supervisor app with recording and registration of behaviour could detect all suspicious events without an impact on the exam performance. Suspicious events were also manually screened after the exam.

The imbalance between registrations for off- or on-campus among the universities is striking but explained by the availability of infrastructure (classes and IT infrastructure). We see indeed that the candidates from larger universities were more likely to register for an off-campus exam. The psychometrics of the exam were comparable between the off- and on-campus group. Here, we can conclude that there was probably no bias induced by the voluntary option for participation off- or on-campus. Other authors compared scores from proctored exams to scores from the traditional exams of the past years and have also found no difference.(9)

The number of technical issues was very low and mainly related to software issues. The complexity of the app (registration of behaviour, recording, etc.) in combination with a multicomponent exam with different question types required very performant IT-equipment. Only in one case a technical issue led to interruption of the exam of the affected candidate. All other issues were solved within an acceptable time span by the supervisors and the developing team. These issues disclosed the vulnerability and low quality of candidates' personal gear. Here, the educational institution should guarantee a fair and safe exam environment for every candidate by offering high quality infrastructure and logistics.

During the entire exam, a team of eight persons was constantly monitoring and in contact with the candidates. Although we did not perform a cost-benefit analysis, we assume that the whole procedure is not cost saving since the off-campus format was run in duplicate to the on-campus exam.

On average, candidates provoked three events of which 'noise/sound' was the most prevalent and critical one. Sound is indeed difficult to avoid or to control (we noticed children crying, birds singing, candidates talking out loud, slamming doors, street works, etc). The developing team immediately increased the threshold sound level during the exam (in the first minutes).

A very small number of candidates was flagged as suspicious by the app but live monitoring during the exam lifted them above suspicion. The supervisors reviewed the recordings of two candidates who were flagged by both the system and the live monitoring during the exam. The records did not reveal any fraudulent behaviour. Prior to the exam, we thoroughly instructed candidates on the registration of suspicious events. This single intervention might have been a decisive argument to withhold candidates from fraud in this high-stake exam. Probably, a simple recording of sound and image combined with a safe web browsing blocking all other webpages might have been as efficient. A less comprehensive

registration of events might reduce the number of technical issues wherefore the supervisors could focus on live monitoring.

During the exams, supervisors limited the number of interventions and warnings to avoid distraction of the candidates. Technical issues were solved, and supervisors sent a warning to only two candidates who were talking to themselves. After the exam, we only received reports of technical issues but no complaints on the interventions or interruptions by the supervisors.

The findings from the exploratory factor analysis indicate that the supervisor app was met with mixed reactions. Considering that candidates took the exam during the COVID-19 pandemic, the supervisor app seems to have provided reassurance that communication with the university was established, in case of technical difficulties. However, the findings indicate that candidates also experienced emotional distress because of the app. Factor 2 identified as emotional distress had high loadings in items that referred to increased stress before and during the exam caused by the supervisor app.

The thematic analysis of the qualitative data provided a more in-depth insight of candidates' perceptions, especially on wellbeing. The third theme relates to the negative impact of the supervisor app on candidates' emotional well-being. A potential explanation for this is that candidates were not sure about what the supervisor app could detect as fraud and suspicious behaviour. Hence, more clear instructions could be beneficial for reducing stress and anxiety while using the app. Nevertheless, the fourth theme comprises candidates' positive emotions and appreciation of the supervisor app. The direct connection with the university through the app seemed to reassure and comfort candidates that assistance was available, in case of technical issues. Lastly, the first and second theme relate more to exam anxiety, and specifically to the format of the exam rather than using the supervisor app. Therefore, they fall outside the scope of this paper, and are not extensively discussed. Regarding technical issues, candidates' comments showed that software issues were the most frequent, confirming what supervisors observed during the exam.

Finally, some recommendations were proposed for improving the supervisor app. The most frequent suggestion was that more clear guidelines were necessary regarding what can be perceived as fraud. Another recommendation was to allow candidates record a video of their surroundings before the exam. These recommendations could be taken into consideration for future improvements of the supervisor app.

## **Strengths and limitations**

The major strength of this intervention is the feasibility to use the app for a large group of candidates for a high-stake exam. Second, we compared the exam outcome of the intervention with a control group composed by candidates who preferred to take the exam on-campus. We cannot completely rule out a bias here but since exam outcomes were comparable for both groups, we consider the risk as low. From

other authors we also know that candidates' preference for computerized or paper based exams does not influence exam outcome.(10)

A limitation of this study is the low number of questions included in the questionnaire. Although the questionnaire could be considered reliable based on the Cronbach's alpha calculation, repeated measurements are necessary to confirm reliability of the scale.

## Conclusion

A sophisticated supervisor app registering behaviour and recording sound and image to prevent fraud during exams has proved to be efficient without affecting the exam outcome. Nevertheless, live monitoring and standing by for solving technical issues induced a high human workload. The supervisor app is also user-friendly, provided that what can be detected as fraud is more thoroughly explained. In future research, a controlled design should compare the cost-benefit balance between the complex intervention of the supervisor app and the combination of the candidates' awareness of being monitored with a safe exam browsing plug in.

## Abbreviations

**AMGP:** Advanced Master's in General Practice

**VUB:** Vrije Universiteit Brussel

**OSCE:** Objective Structured Clinical Examination

**GP:** General Practitioner

**API:** Application Programming Interface

**SEB:** Safe Exam Browser

**ENT:** Otolaryngology

**KMO:** Kaiser-Meyer-Olkin

## Declarations

## Ethical approval

According to the ruling policy and regulations at the involved universities (KU Leuven, University of Ghent, University of Anwerp, VUB), no participants' informed consent was needed in order to report the outcome measures of the supervisor app, because we used anonymized secondary data. Nevertheless, participants (students) were asked to consent with the ongoing procedure of the exam and to contribute

to the rollout of it by the Interuniversity Center for GP Training that initially collected the data for evaluation purposes.

For the post-test survey, ethical approval was granted by the Social and Societal Ethics Committee of the KU Leuven with the following approval number: *G-2020-2262-R2(MAR)*.

## **Consent for publication**

All participants have been informed and given consent for publication.

## **Competing interests**

None.

## **Availability of data and materials**

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

## **Authors' contributions**

VA designed the questionnaire, collected, and analyzed the data, wrote, and revised the manuscript. BS set up the test setting, collected and analyzed the data and wrote and revised the manuscript. SP and JE critically reviewed and revised the manuscript. JW set up the test setting, reviewed and approved the manuscript. All authors have read and approved the manuscript.

## **Funding**

None.

## **Acknowledgements**

We would like to thank and acknowledge An Stockmans, Guy Gielis, Bert Aertgeerts, Roy Remmen, Dirk Devroey, Aram Hovsepyan, Chavdar Gushterov, Peter Decat, Madieke Maseman for their intellectual and logistic support in the developing and unrolling of this plug in.

## **Authors' information**

[1] KU Leuven, Academic Center for General Practice, Department of Public Health and Primary Care, 3000 Leuven, Belgium

<sup>2</sup> Evidence Based Practice, EBMPracticeNet, 3000 Leuven, Belgium

<sup>3</sup> Department General Practice, Monash University, 3800 Melbourne, Australia

<sup>4</sup> KU Leuven, Department of Cellular and Molecular Medicine, 3000 Leuven, Belgium

<sup>5</sup> University of Antwerp, Center for General Practice/Family Medicine, Department of Primary and Interdisciplinary Care, 2610 Wilrijk, Belgium

## References

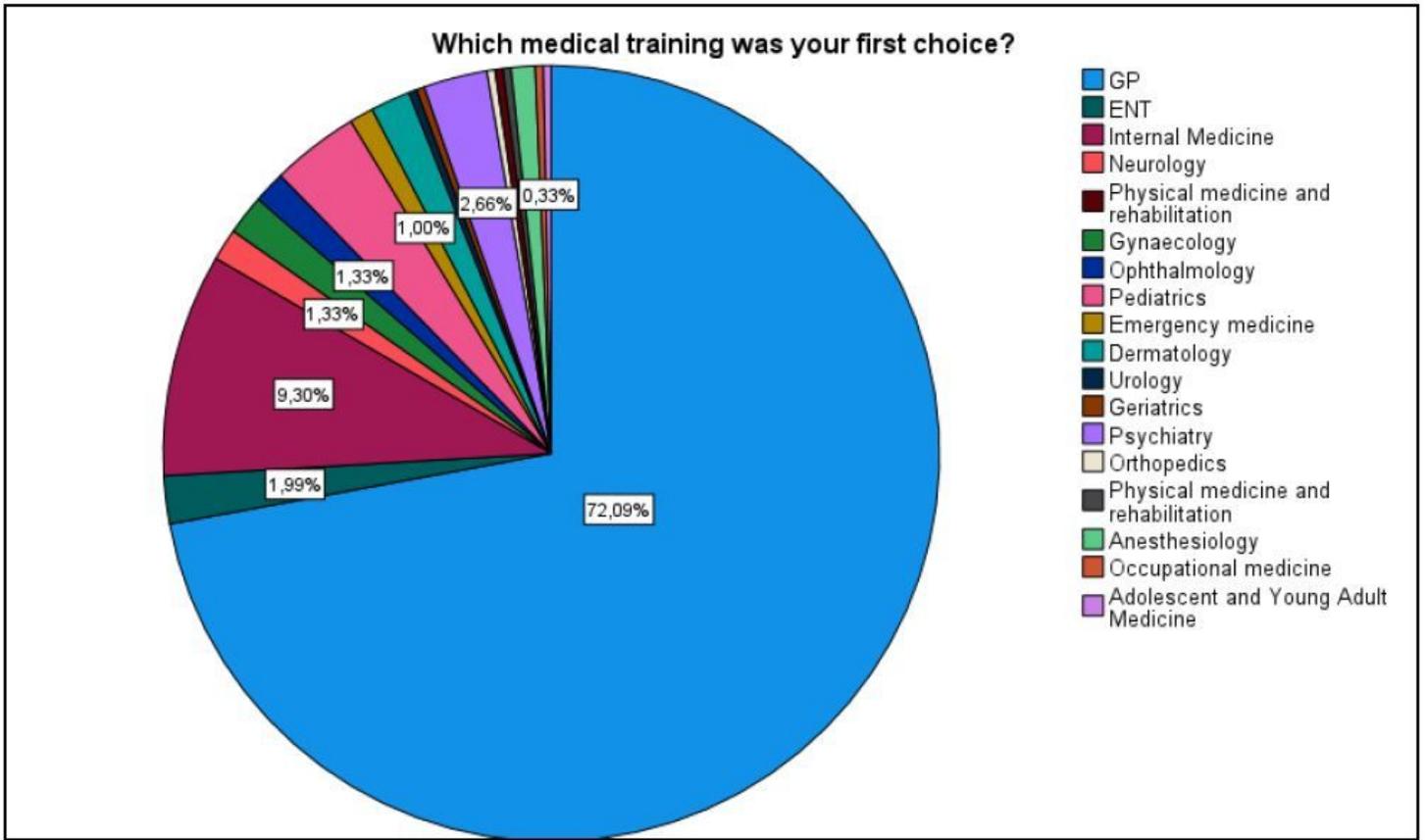
1. Ogilvie RW, Trusk TC, Blue AV. Students' attitudes towards computer testing in a basic science course. *Med Educ.* 1999;33(11):828-31.
2. Andreou V, Eggermont J, Gielis G, et al. Proficiency testing for identifying underperforming students before postgraduate education: a longitudinal study. *BMC Medical Education.* 2020;20(1):261.
3. Schoenmakers B, Wens J. Proficiency testing for admission to the postgraduate family medicine education. *J Family Med Prim Care.* 2018;7(1):58-63.
4. Munshi F, Alsughayyer A, Alhaidar S, et al. An Online Clinical Exam for Fellowship Certification during COVID-19 Pandemic. *Med Educ.* 2020.
5. Field A. *Discovering Statistics Using SPSS*: SAGE Publications; 2009.
6. Costello AB, Osborne J. Best Practices in Exploratory Factor Analysis: Four Recommendations for Getting the Most From Your Analysis. *Practical Assessment, Research & Evaluation.* 2005;10:1-9.
7. Tavakol M, Dennick R. Making sense of Cronbach's alpha. *Int J Med Educ.* 2011;2:53-5.
8. Clarke V, Braun V. Thematic analysis. *The Journal of Positive Psychology.* 2016;12:1-2.
9. Sam AH, Reid MD, Amin A. High-Stakes Remote-Access Open-Book Examinations. *Med Educ.*n/a(n/a).
10. Hochlehnert A, Brass K, Moeltner A, et al. Does medical students' preference of test format (computer-based vs. paper-based) have an influence on performance? *BMC Med Educ.* 2011;11:89.

## Figures

1. **Blur page** - Suspicious
  - The user has switched to another browser tab or other program
2. **Focus page**- Suspicious
  - The user had returned his focus to the page. It can be used after Blur to determine how long the user had been gone
3. **Close page**- Suspicious
  - The user has closed the current tab or has opened different page in the same tab
4. **Load page**
  - When the user loads the page. It can be used after Close to determine how long the user had been gone
5. **Keyboard events (basic key logger)**
6. **Connected to internet**
7. **Disconnected from internet** - Suspicious
  - The user can stop the internet to stop the video recording. It will be sent to the server when the internet is connected
8. **Mouse click (including click location coordinates)**
9. **Sound detected** - Suspicious
  - This tracks any noise above certain volume level and with length > 1ms
10. **Admin started watching** (this is marked in blue in the tracking page)
11. **Admin started talking** (this is marked in blue in the tracking page)
12. **Finalize component** (this is needed because when the user completes the component then Close is triggered)

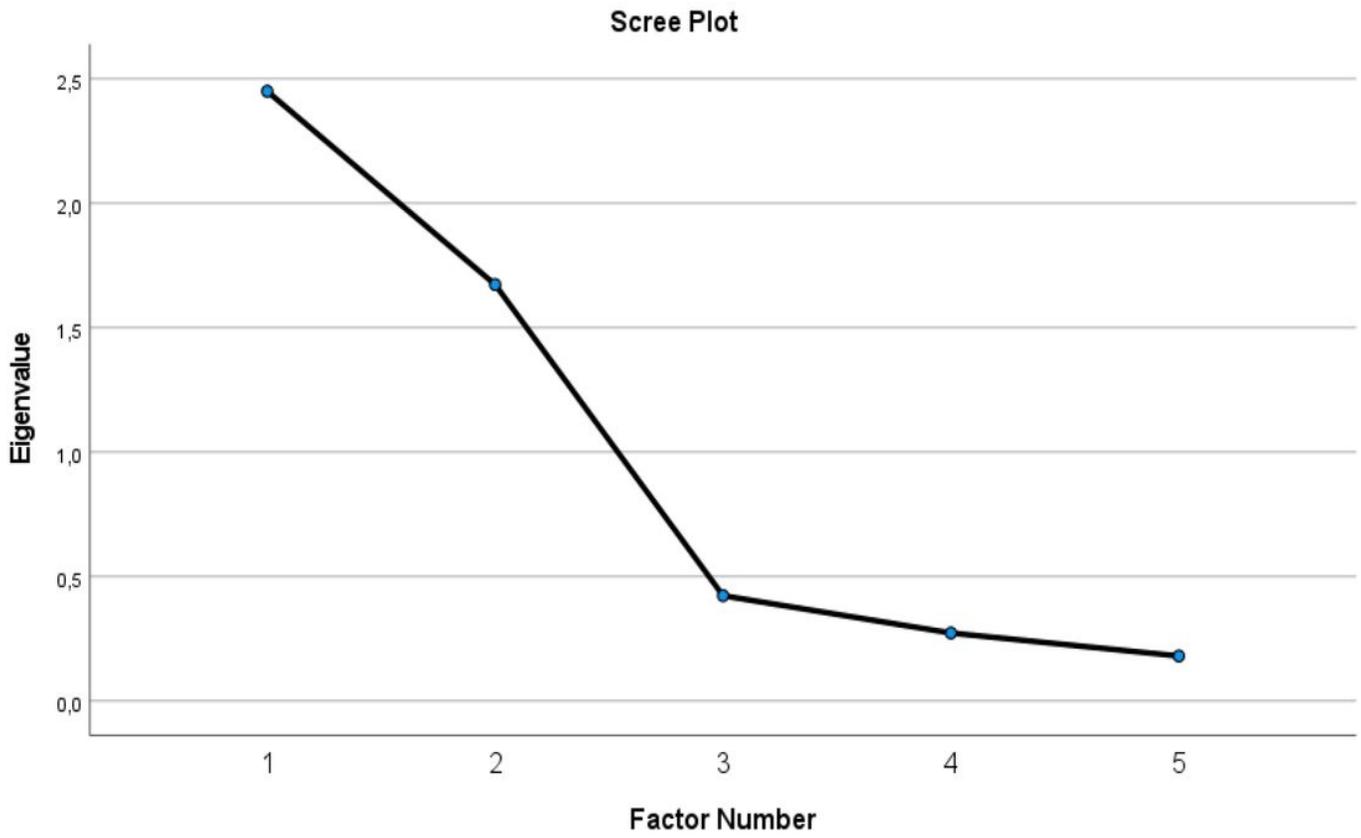
**Figure 1**

Type of events tracked and traced by the supervisor app



**Figure 2**

Pie chart of candidates' first choice of medical training



**Figure 3**

Scree plot of the eigenvalues of the factors