RAYUELA
a fun way to fight cybercrime

Deliverable Report

D6.6 Open: Key conclusions/risk factors on the main human factors for cybercrime in children and young adults

Document Information and contributors

Deliverable No.	D6.6	Work Package No.	WP6	Task/s No.	Task 6.4
Work Package Title	DATA ANALYSIS AND INTERPRETATION ON PROFILES FROM POTENTIAL YOUNG VICTIMS AND OFFENDERS				
Linked Task/s Title	T6.4 interpretation of collected data on the impact of human factors on young cybercrime victims and offenders				
Status	Final	(Draft/Draft Final/Final)			
Dissemination level	PU	(PU-Public, PP, RE-Restricted, CO-Confidential)			
Due date deliverable	30/09/2023	Submission date		18/10/2023	
Deliverable version	1.0				
Deliverable responsible	COMILLAS				

Contributors	Organisation	Reviewers	Organisation
Jaime Pérez	COMILLAS	Violeta Vazquez	ZABALA
Gabriel Valverde	COMILLAS	Irene Serrano	ZABALA
Mario Castro	COMILLAS	Lokesh Sharma	NEC
Gregorio López	COMILLAS	Jelle Janssens	UGENT
María Reneses	COMILLAS		
María Riberas	COMILLAS		

(*) Note on the contributors to this deliverable: Since the RAYUELA project deals with sensitive data, the data-driven results and conclusions coming from task 6.2 and 6.3 needs to be reviewed and interpreted from a multidisciplinary perspective by experts. Thus, the interpretation of the obtained results included in this deliverable has been carried out with the support and participation of experts within the consortium coming from social sciences (BPI, TARTU, UGENT), education (EA, UCLL), and LEAs (PLV, PJ, EPBG, PSNI), as well as with feedback from all the partners in the consortium through the organised workshop.

Document History

Version	Date	Comment
0.1	01/09/2023	1st draft – Definition of structure and content
0.2	12/09/2023	First complete draft
0.3	30/09/2023	Improved draft ready for review
1.0	15/10/2023	Final version including reviewers' comments

TABLE OF CONTENTS

List of Abbreviations

Abbreviation	Description
CB	Cyberbullying
CH	Cyber Harassment
BN	Bayesian Network
DAG	Direct Acyclic Graph
WP	Work Package

Executive summary

The fourth task of Work Package 6 in the RAYUELA project aims to clarify and refine the results obtained for the cybercrimes under consideration (affecting minors). Given the sensitivity of the subject matter, it is crucial to review and interpret the data-driven results and conclusions from the project through a multidisciplinary lens of expertise, including law enforcement agencies, psychologists, sociologists, educators, and more. This holistic approach is essential to clarify and refine the knowledge acquired.

In this document, we combine the quantitative results obtained applying causality and Bayesian statistics to analyse the data gathered through the videogame and the results of previous work packages addressing human and technological factors. All these results are critically discussed with the other members of RAYUELA. This discussion occurred during an interactive workshop held in Zagreb (Croatia) on September 28 2023.

The main objective of this document is to clarify the impact of the analysed human and technological factors in cybercrime affecting minors. Subsequently, the document also aim to provide recommendations and guidelines that may be useful for law enforcement agencies and policymakers. Our results suggest that considering the available data and technical limitations of the methodologies employed:

- The RAYUELA serious game represents a **helpful social science research tool** to study the cybercrimes under consideration.
- **The data obtained through the serious game are relevant to explaining and predicting** the risk of suffering/committing cybercrimes.
- **Demographic variables or those obtained through psychological tests do not have significant relevance** when considered independently in any of the cybercrimes studied except in the case of cyberbullying offences, especially the variable indicating previous victimisation, which strongly influences the experiments.
- **Future research might also include risk taking behaviour to contribute to the overall findings.**
- **Some of the methodologies** used in the experiments to "interrogate" the model once it is trained **are non-standard** since we have had to adapt them to the specific needs of our problem. This implies that some results and conclusions could be distorted.

These findings confirm the suitability of our chosen methodology, not just for predictive purposes but also for its prescriptive capability, which has been validated with expert input. However, it is imperative to reiterate the methodological limitations, emphasising the need to interpret these results with caution. We hope the insights gained through this analysis will be valuable in advancing our understanding of young cybercriminality and developing public policy safeguard minors from engaging in inappropriate online activities and improve their online experience.

1. Introduction

The primary goal of this document is to collect the results and conclusions drawn from the statistical analyses and computational modelling carried out in Tasks 6.3 and compare them with the results obtained in the previous research conducted in WP1 and WP2. Once these results have been collected and compared, they are discussed with the rest of the RAYUELA members through an interactive workshop to clarify and refine the conclusions drawn and the knowledge acquired. Workshop participants included Law Enforcement Agencies (LEAs), psychologists, sociologists, criminologists, educators and engineers.

In Task 6.3, **causality** and **Bayesian Networks** (BN) have been used. Our methodology (that we refer to as Probabilistic Causal graphs) aims to understand how one event or variable causes another to occur, understood as "intervention". Namely, how if we intervene in one variable expert the outcome to change. The second ingredient is that this assessment of the degree of change is done probabilistically, so the methodology also provides a measure of relevance and uncertainty in the estimation. Interestingly, causal inference methods estimate the causal effect of one variable on another from observational data while controlling for confounding variables that may influence the relationship. As discussed in Task 6.3, the mathematical tool behind this method is the Bayesian Network representing the dependency structure of a set of variables and their joint probability distributions.

A causality-based approach is convenient when dealing with sensitive research topics, such as cybercrime in minors. We wish to understand such issues better to optimise prevention and mitigation strategies. These techniques natively support missing data, where other statistical or modelling methods can struggle. It allows us to identify the strength of causal relationships and control for confounding variables, which can remove biases in relationships between the independent and dependent variables. A side benefit of the graphical approach is that it forces us to make explicit assumptions and hypotheses, leading to an open, transparent and critical debate.

Besides the analysis performed in WP6, in this deliverable, we also include the analysis obtained from a live discussion on the expectations and opinions of the other members of the RAYUELA consortium through an interactive workshop. In a joint live work session (workshop format), we presented the results obtained from the data analysis. We compared these results with those from WP1 (and WP2 in the case of technological cyber threats). The workshop dynamics consisted of presenting the results of a real-time questionnaire. The questions were collected and shared in real time using the tool Wooclap. Through these dynamics, we sought to validate our results, resolve discrepancies, and obtain valuable qualitative insights that complement our quantitative research.

The rest of the document is organised as follows: in Sec. 2, we explain in more detail the methodology used to collect the beliefs and discussions of the other members of RAYUELA about the cybercrimes under consideration. In Sec. 3, we summarise the results obtained in the previous WP6 tasks and present the main results of the qualitative and quantitative analysis of the questionnaire and the discussion. Finally, in Sec. 4, we present the conclusions drawn from all the work.

2. Methodology

This section outlines the methodology employed to gather the opinions and discussions of other RAYUELA members regarding the cybercrimes under consideration, the results obtained in Task 6.3 and their comparison with the results from previous work packages (namely, WP1 and WP2).

As mentioned above, this opinion gathering, and discussion took place in a workshop with the members of RAYUELA. In this workshop, participants were asked in an interactive questionnaire about their opinions on potential risk factors/profiles identified, proposals for intervention, discussions on disagreements between data analysis and insights from the research carried out in WP1 and WP2, and finally, a series of conclusions/highlights were collected for their opinion on how much they agreed with those statements. The aim was also to demonstrate to participants the potential impact of interventions on each cybercrime, discuss their effectiveness and address various points of view on which factors are the most influential and how to intervene on them, all around data.

To carry out the online questionnaire and to be able to visualise the results in real-time, which facilitates the discussion, we used the Wooclap2 tool. Conversations and discussions were recorded and transcribed, and they are an integral part of the conclusions.

It should be noted that a significant part of the questionnaire and discussion was devoted to cyberbullying (CB) since this cybercrime is the only one for which we also have a validated psychological questionnaire that the minors participating in the RAYUELA pilots had to fill out. For CB, we have a measure like a "ground truth". Furthermore, experts agreed that it was the most possible cybercrime to measure. However, the discussion was focused on the rest of cybercrimes in any case.

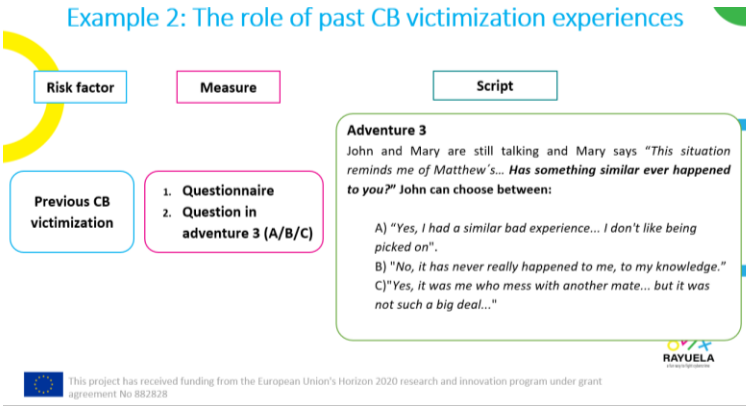
The **agenda** of the workshop was as follows:

1. Introduction
    a. Main goals of the workshop
    b. First examples of the tool Wooclap
2. Risk patterning (Summary of the concept of "risk pattern" (combinations of variables) and examples.)
3. Prescription
    a. Effectiveness vs. Easiness
    b. Examples of network intervention simulation
4. Consensus and Discussion
5. Highlights and Discussion

Through the RAYUELA pilots conducted in schools with minors, we collected 1794 play sessions. Participants were between 12 and 16 years old (Mean=14.05, SD=1.38), where 57% identified themselves as males, 44% as females and 1% as non-binary. In Annex 1, we present an exploratory data analysis of all collected data. The participants must register on a website where they will receive a user ID and password to enter the game. The **variables** collected from the participants during these pilots (anonymously and optionally) are the following:

i) Demographics: Age, gender, sexual orientation and migratory background.
ii) Technological: Daily hours spent on the Internet (for leisure).
iii) Psychological & Sociological Questionnaires:
   a) A short version of the Big Five Inventory of personality traits questionnaire [1], where a value from 1 to 10 is inferred for each of these traits (Agreeableness, neuroticism, extraversion, conscientiousness, openness to experience)
   b) The Multidimensional Scale of Perceived Social Support [3] estimates social support from friends, significant others and family. These supporting dimensions are classified into 3 categories: low, medium and high.
   c) Rosenberg self-esteem scale [2] infers the user's self-esteem and classifies it into three categories: low, medium, and high.
   d) Questionnaire on cybervictimization and cyberaggression: European Cyberbullying Intervention Project Questionnaire [4]
iv) Gameplay (inside RAYUELA's serious game):
   a) Answer to game decisions: The minors proceed to play the RAYUELA game, where they first personalise an avatar and then enter the adventures. In these adventures, they must decide what the character should do. The decisions made and the response times are stored for later analysis.
   b) Honesty: Final question about differences between their behaviour in the game and in reality («Do you think the answers you have given in the game are similar to how you would act in real life? »).

When designing the methodology for collecting data from the children participating in the pilots, the research carried out by WP1 and WP2 on the cybercrimes under study was carefully considered. Moreover, some of the indicators/factors to be studied were measured both in the questionnaires and in the RAYUELA serious game itself. Fig. 1 shows an example of the risk factor of previous CB victimisation. This methodology, combined with the final question on the honesty of the answers given, can be used to **calibrate** estimates in statistical analyses and modelling.



Example 2: The role of past CB victimization experiences

Risk factor	Measure	Script
Previous CB victimization	1. Questionnaire 2. Question in adventure 3 (A/B/C)	**Adventure 3** John and Mary are still talking and Mary says *"This situation reminds me of Matthew´s...* ***Has something similar ever happened to you?"*** **John can choose between:** A) *"Yes, I had a similar bad experience... I don't like being picked on".* B) *"No, it has never really happened to me, to my knowledge."* C)*"Yes, it was me who mess with another mate... but it was not such a big deal..."*

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 882828

Figure 1. Slide from the workshop showing an example of how the game adventure was designed to estimate the risk factor 'Prior CB victimisation'. This factor has been evaluated independently using a validated questionnaire that players must fill in and in a question within the RAYUELA serious game itself.

3. Results

3.1 Description of the workshop participants

During the Consortium Meeting held in Zagreb, we held a workshop with 29 participants, distributed per institution, as shown in Fig. 2.

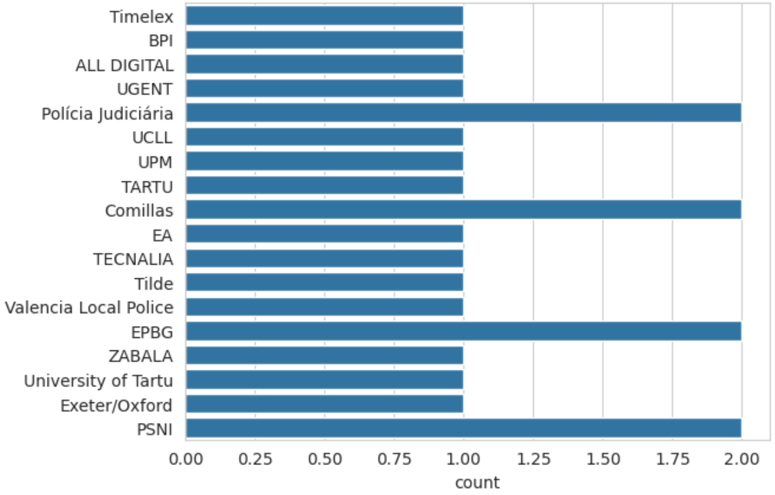


Figure 2. Distribution of participants per institution in the workshop held during the 7th RAYUELA consortium meeting held in Zagreb on September 28, 2023.

Participants came from various backgrounds, including engineers, psychologists, and LEA members. This diverse experience brought a wide range of knowledge about the cybercrimes we discussed. However, involvement in the RAYUELA project led to responses toward options 1 and 2, indicating a modest level of understanding. In a previous RAYUELA meeting held in Valencia, March 2023), we also asked the participants about their prior knowledge. In Fig. 3-5, we summarise the distribution of “prior knowledge” or “expertise” for Cyberbullying, Online Grooming and Cybersecurity from the participants in Valencia’s meeting.

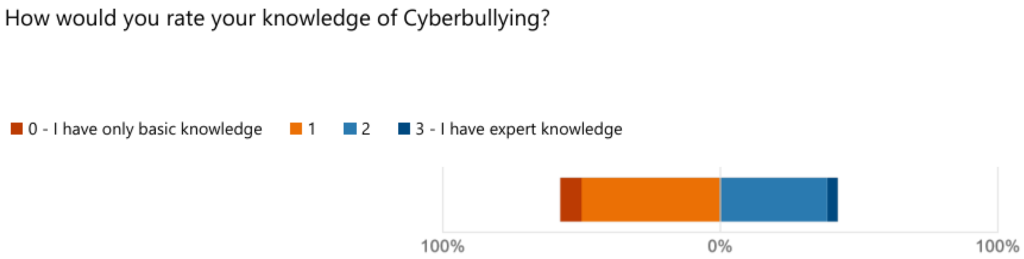


Figure 3. Distribution of self-perceived expertise for different Cyberbullying

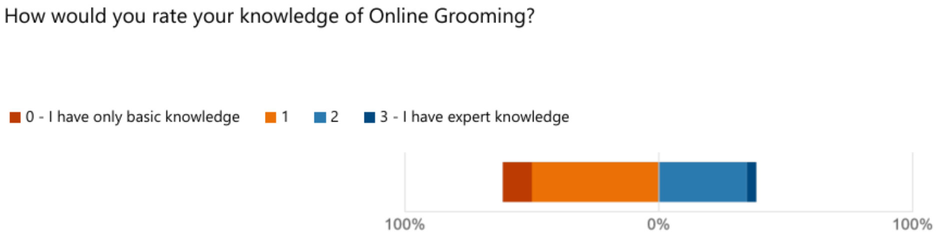


Figure 4. Distribution of self-perceived expertise for different Online Grooming

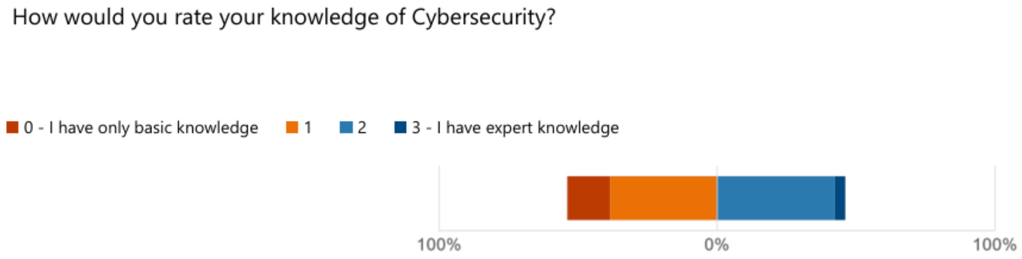


Figure 5. Distribution of self-perceived expertise for Cybersecurity.

How accurately do you expect we will be able to measure the following roles in cybercrime?

As discussed above, in the RAYUELA pilots, we collected sociodemographic information, in-game responses and a set of validated questionnaires from the players. Some questionnaires were presented to the players before and some after to avoid biasing their attitudes during the game.

One of the main goals of this workshop was to calibrate the participants' expectations of the game outputs, not only to fulfil its function as an educational resource to prevent risky attitudes related to cybercrime but also to deploy a scientific tool to larger populations than those traditionally accessible to a questionnaire or interview-based methodologies. However, this second goal has yet to be previously demonstrated (being this is one of the strengths of RAYUELA), and we wanted to determine the attitudes and expectations of other members of RAYUELA (not working in WP6) in this regard. Hereafter, we compare their responses to 5 questions related to those expectations. In Fig. 6, the results of this question in the questionnaire are shown. The results seem to indicate that workshop participants have good expectations, in all the considered cybercrimes, about the ability of the game to measure potential victims/aggressors.

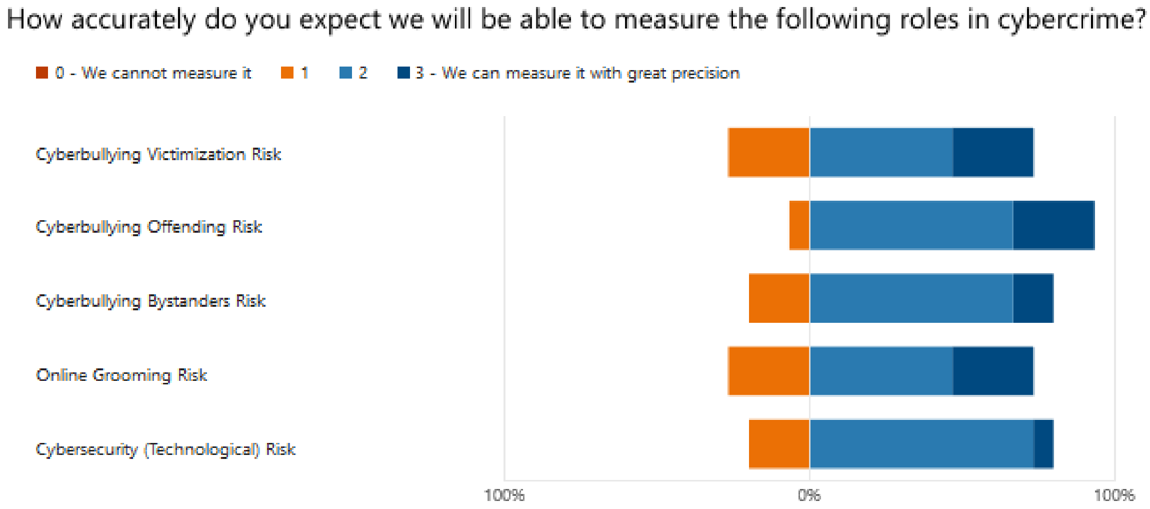


Figure 6. Distribution of the participants' expectations on the potential ability of the RAYUELA game to measure the cybercrimes under consideration.

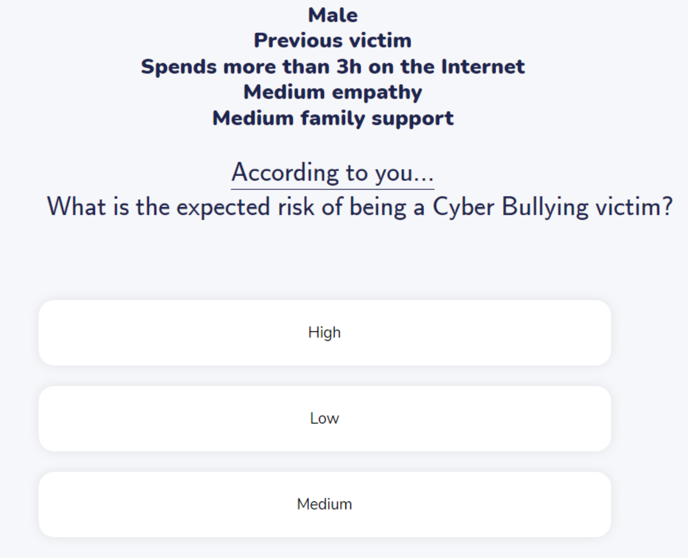
3.2 Description of the workshop questions

We can classify the workshop questions into four categories:

I. Risk Patterning

In this category, we focus on CB profiles since, as we have commented above, this cybercrime is the only one for which we also have a validated psychological questionnaire that the minors participating in the RAYUELA pilots had to fill in. Namely, we have a measure like a "ground truth" for CB.

In collaboration with the psychologists from WP1, we selected only five sociodemographic variables that we know to be crucial and commonly used in their daily work with adolescents. Participants are presented with a profile with specific demographic/personal characteristics and are asked to rank the a priori risk this profile would have (based on their opinion/intuition/expert knowledge). An example of these questions is shown in Fig. 7.



Male
Previous victim
Spends more than 3h on the Internet
Medium empathy
Medium family support

According to you...
What is the expected risk of being a Cyber Bullying victim?

High

Low

Medium

Figure 7. Example of a Risk Patterning question in the Wooclap tool. Participants are asked about their opinion/intuition about the a priori risk of a profile with specific demographic/personal characteristics to suffer cyberbullying (in this case).

We established two profiles representing high, neutral, and low risk of being cyberbullies, and low, neutral, and high risk of being victims—both based on real-world occurrences and the probabilities generated by the models. Participants provided their assessments of whether they considered these profiles as more or less risky in each case. The effects of these variables on the model's predictions, using responses from various questions within the video game, are also presented. Before this, experts were asked how these variables might affect the predictions to validate if their intuitions aligned with the model's conclusions regarding the impact of these responses. Fig. 8 and 9 show an example of a risk patterning profile and the interface to gather data from the workshop participants.

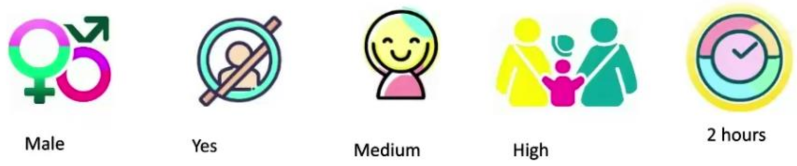
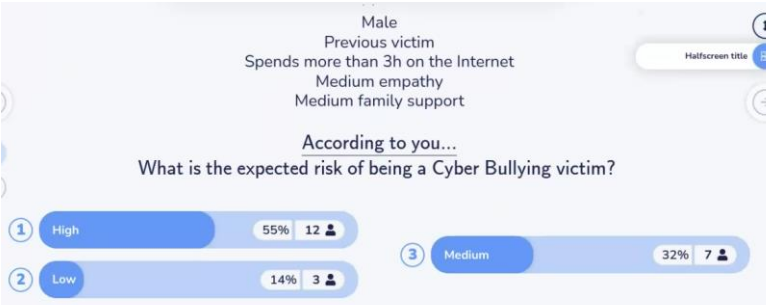


Figure 8. Example of risk pattering in the model developed within RAYUELA.



Male
Previous victim
Spends more than 3h on the Internet
Medium empathy
Medium family support

According to you...
What is the expected risk of being a Cyber Bullying victim?

1	High	55%	12
2	Low	14%	3
3	Medium	32%	7

Figure 9. Example of results gathered from the participants (Risk patterning).

II. Effectiveness vs. Easiness of interventions on the available variables

Participants are shown the list of variables with which the project has been working, therefore considered important and measurable. From an intervention perspective, participants are then asked to indicate how effective an intervention on each variable would be and subsequently how easy it would be to carry out (considering technical, logistical or ethical constraints).

At this point of the workshop, attendees were shown a simulation of interventions (and marginalisation) using the Bayesian network-based model and the GeNIe[1] software tool. This example was intended to illustrate the results obtained based on the data collected in the pilots. Fig. 10 shows a screenshot of the GeNIe software.

[1] https://www.bayesfusion.com/genie/

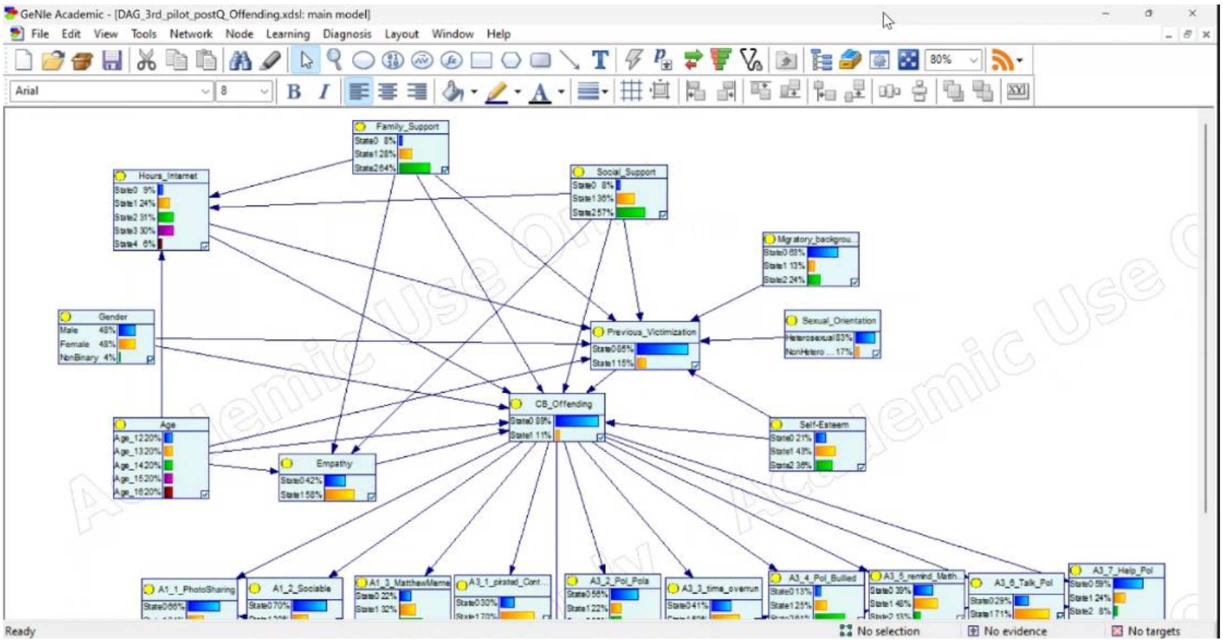
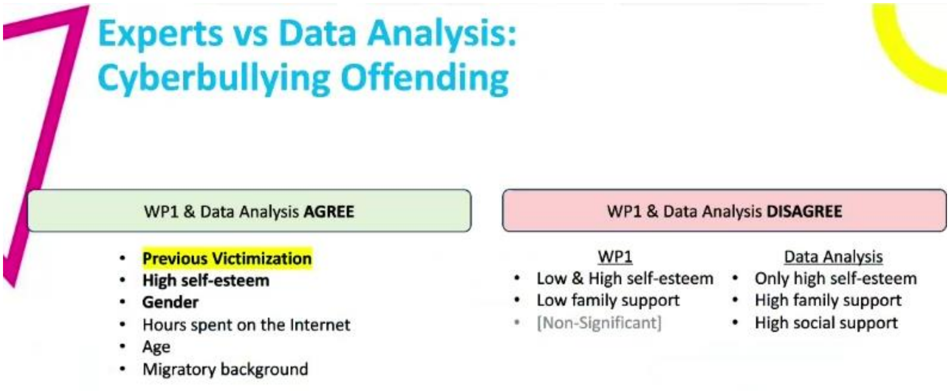


Figure 10. Screenshot of the GeNIe software to simulate interventions/marginalisation in the Bayesian Networks based model.

III. Consensus and Discussion: Highlights

In this workshop phase, we have compared the results from WP6 (data analysis and modelling) with those from WP1 (literature review, child surveys, judgment analysis and interviews) and, in the case of technological cyberthreats, with WP2. We have highlighted areas of consensus between WP1 and WP6, as well as where discrepancies exist. This process has led to debate and discussion, especially regarding the disagreements, seeking possible explanations. Fig. 11 shows an example of this comparison.

Subsequently, participants were presented with a series of possible conclusions or highlights for each cybercrime, where they were asked to indicate how much they agreed/disagreed with these statements. Fig. 12 shows an example of this part of the workshop.

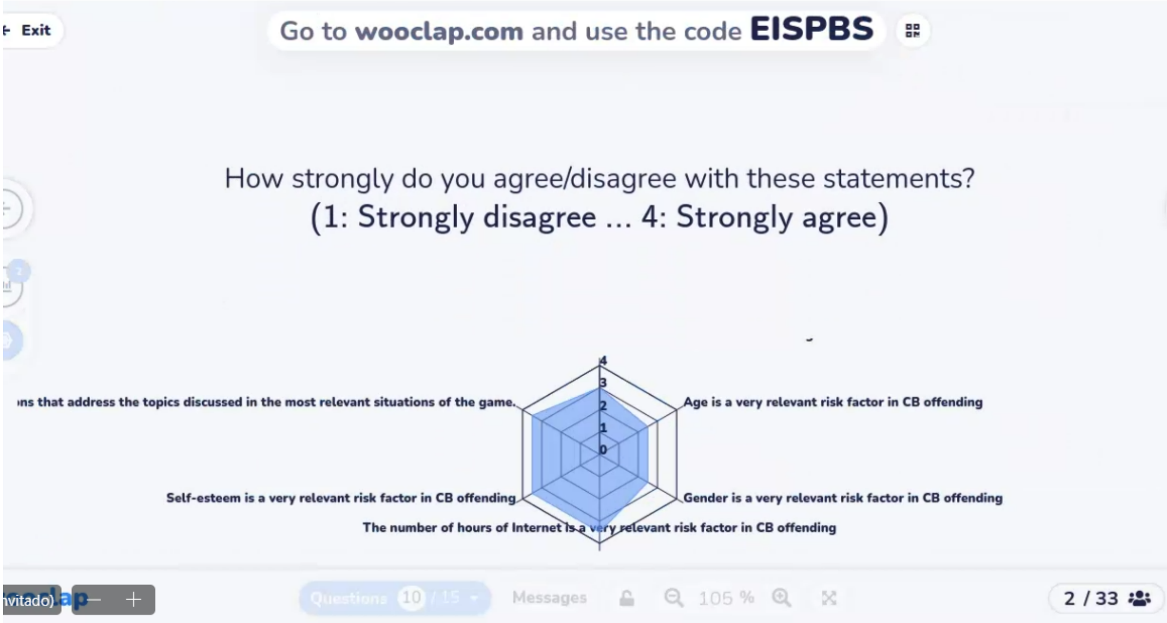


Experts vs Data Analysis: Cyberbullying Offending

WP1 & Data Analysis **AGREE**
• **Previous Victimization** • **High self-esteem** • **Gender** • Hours spent on the Internet • Age • Migratory background

WP1 & Data Analysis **DISAGREE**	
WP1	Data Analysis
• Low & High self-esteem	• Only high self-esteem
• Low family support	• High family support
• [Non-Significant]	• High social support

Figure 11. Example of agreement/disagreements between WP1 and WP6 results in cyberbullying offending.



Go to **wooclap.com** and use the code **EISPBS**

How strongly do you agree/disagree with these statements?
(1: Strongly disagree … 4: Strongly agree)

Exit
4
3
2
1
0
ns that address the topics discussed in the most relevant situations of the game.
Age is a very relevant risk factor in CB offending
Self-esteem is a very relevant risk factor in CB offending
Gender is a very relevant risk factor in CB offending
The number of hours of Internet is a very relevant risk factor in CB offending
nvitado)
Questions 10 / 15
Messages
105 %
2 / 33

Figure 12. Example of possible conclusion/highlights presented to the participants in order to gather their degree of agreement/disagreement with each statement.

3.3 Results by Cybercrime

This subsection summarises the most relevant results from Tasks 6.3. The main objective was to apply statistical analysis and computational modelling techniques to the data collected in the RAYUELA pilots to conclude the most relevant risk factors/profiles for the considered cybercrimes.

3.3.1 Cyberbullying

3.3.1.1 Cyberbullying: Summary of results from the data analysis

The results below summarise those obtained in deliverable D6.4, which is also openly available for anyone wishing to consult it.

- **Research Question 1**: Which variables are most strongly related to the risk of suffering/committing cyberbullying?

Based on the selected causal structure, available data, and metrics, the most relevant variables of having committed and suffered CB are shown in Table 1. The results of studying the influence of the variables separately (i.e., one at a time) seem to indicate that the game and the decisions that players must make are very relevant in explaining the CB Offending and Victimisation variables. It also seems that explaining (and therefore predicting) the variable CB Offending seems easier from the available data. The only relevant variable directly related to demographics or psychological tests is "*Previous CB Victimization*", in the case of explaining the aggression variable. This result has also been found in the literature and WP1 research.

Table 1. Most strongly related variables for cyberbullying offending and victimisation. This analysis studies variables one at a time within the Bayesian network.

CB Offending	CB Victimisation
Adventure 3 Question 3: Time Overrun	*Adventure 3 Question 5: Remind Matthew*
Adventure 3 Question 4: Pol Bullied	*Adventure 3 Question 3: Time Overrun*
Adventure 3 Question 5: Remind Matthew	*Adventure 3 Question 4: Pol Bullied*
Adventure 1 Question 3: Matthew Meme	
Previous CB Victimisation	
Adventure 3 Question 1: Pirated Content	
Adventure 3 Question 2: Pol Pola	

- **Research Question 2:** What combinations of variables make it possible to construct meaningful risk profiles for suffering/committing cyberbullying?

Based on the selected causal structure, available data, and metrics, the most relevant characteristics shared by risk profiles of having **committed CB** are shown in Table 2. Equivalently, they are shown in Table 3 for the case of CB victimisation. The results of studying the influence of demographic and personal variables (excluding game questions variables) in combination indicate that 'previous CB victimisation' remains crucial in explaining aggression. By combining the variables shown in the tables, we can create a typical aggressor or victim profile. However, it should be noted that this profile is by no means determinant. In other words, there is no single profile of either perpetrator or victim, so these results should be interpreted cautiously. We can also see from the figures in the tables that when using combinations of variables, the difference in importance between game and demographic/personal variables is reduced, and they become very similar in terms of "profiling effectiveness".

Table 2. Summary of information obtained from the multifactorial analysis of various risk factors examined for Cyberbullying offenders. The colour scheme in the graph denotes variable importance, with green representing high prevalence, yellow for moderate prevalence, and red for somewhat prevalent.

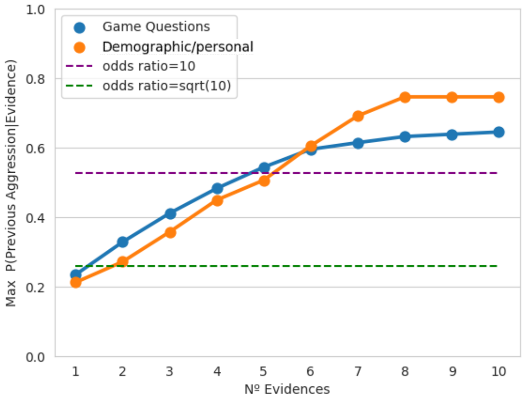
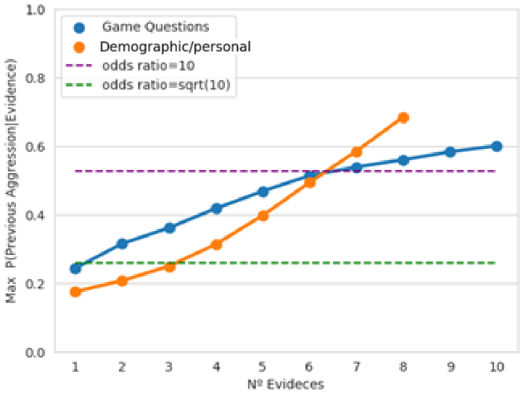
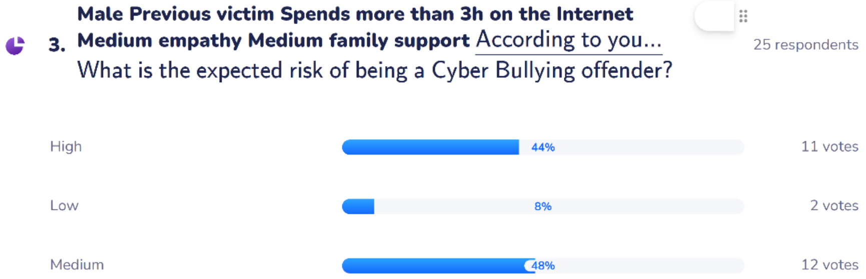
Cyberbullying Offender	
Most prevalent risk factors/variables	**Evolution of "profiling effectiveness" as a function of the number of variables/evidences used**
Previous CB Victimization	
Gender Male	
High Self-esteem	
High family support	
High social support	
Medium-high daily hour Internet (3-4h)	
Age = 16	
Heterosexual	
Age = 14	
Medium daily hours Internet (2-3h)	

Table 3. Summary of information obtained from the multifactorial analysis of various risk factors examined for Cyberbullying Victimization. The colour scheme in the graph denotes variable importance, with green representing *high prevalence*, yellow for *moderate*, and red for *somewhat prevalent*. In the right panel, the dashed green and purple lines denote the Bayes factor levels of good and strong evidence, respectively.

Cyberbullying Victimisation	
Most prevalent risk factors/variables	**Evolution of "profiling effectiveness" as a function of the number of variables/evidences used**
Heterosexual	
No migratory background	
High social support	
Male	
Medium family support	
Medium self-esteem	
High family support	
Age = 15	
Female	
Medium-high daily hour Internet (3-4h)	

3.3.1.2 Cyberbullying: Summary of results from the workshop

Below are the workshop participants' answers and the aggregate results for each. Fig. 13-17 show the questions related to CB **Risk Patterning**, in which participants were shown an example of a perpetrator or victim profile and had to estimate the level of associated risk they considered based on their experience or expert knowledge. According to the data analysis and modelling performed by WP6, Fig. 13 and 14 show high-risk profiles, Fig. 15 medium risk and Fig. 16 low risk. In addition, Fig. 17 adds to the profiling of the answer to a question within the video game, which makes it possible to change its estimated risk from high to low radically. Generally, there seems to be an agreement between the data analysis results and the workshop participants' opinions.

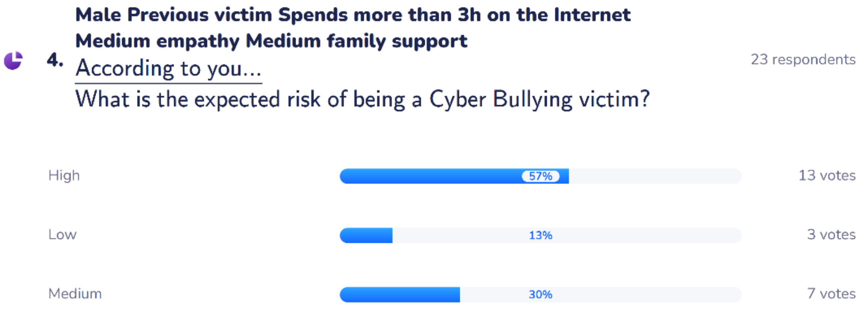


3. **Male Previous victim Spends more than 3h on the Internet Medium empathy Medium family support** According to you...
What is the expected risk of being a Cyber Bullying offender?

25 respondents

High	44%	11 votes
Low	8%	2 votes
Medium	48%	12 votes

Figure 13. First CB risk patterning question. Modelling and data analysis indicated this profile as a high-risk probability. Responses can be Low, Medium, or High risk.

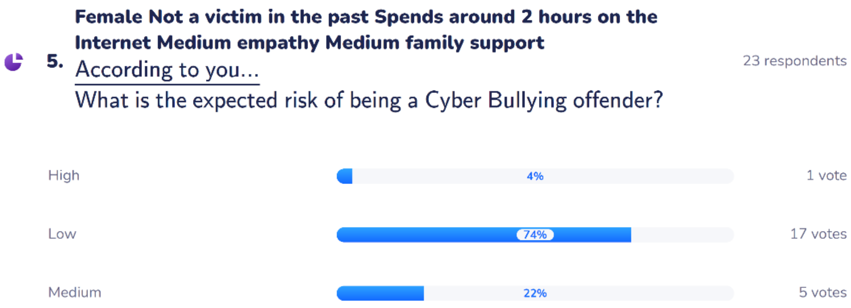


4. **Male Previous victim Spends more than 3h on the Internet Medium empathy Medium family support**
According to you...
What is the expected risk of being a Cyber Bullying victim?

23 respondents

High	57%	13 votes
Low	13%	3 votes
Medium	30%	7 votes

Figure 14. Second CB risk patterning question. Modelling and data analysis indicated this profile as a high-risk probability. Responses can be Low, Medium, or High risk.

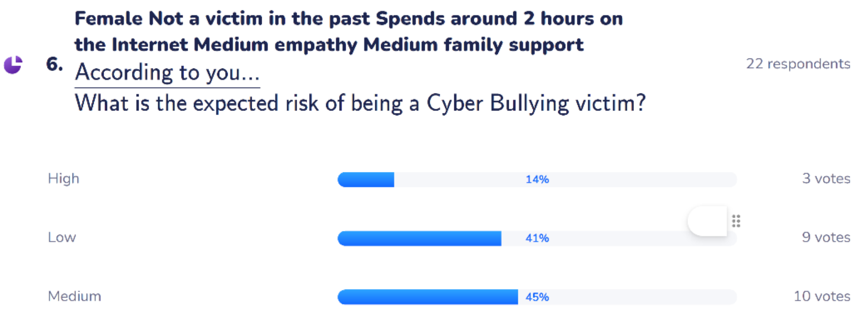


5. **Female Not a victim in the past Spends around 2 hours on the Internet Medium empathy Medium family support**
According to you...
What is the expected risk of being a Cyber Bullying offender?

23 respondents

High	4%	1 vote
Low	74%	17 votes
Medium	22%	5 votes

Figure 15. Third CB risk patterning question. Modelling and data analysis indicated this profile as medium risk probability. Responses can be Low, Medium, or High risk.

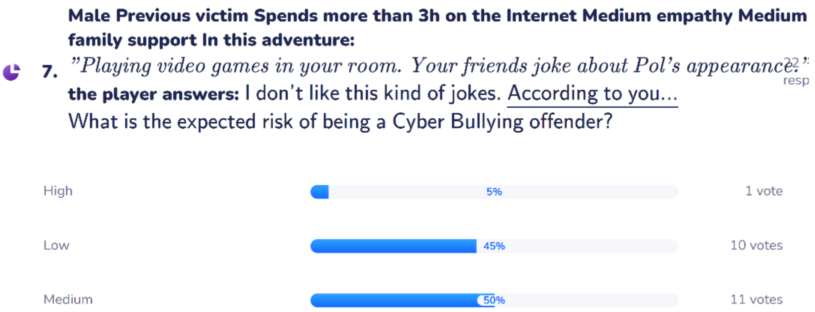


6. **Female Not a victim in the past Spends around 2 hours on the Internet Medium empathy Medium family support**
According to you...
What is the expected risk of being a Cyber Bullying victim?

22 respondents

High	14%	3 votes
Low	41%	9 votes
Medium	45%	10 votes

Figure 16. Forth CB risk patterning question. Modelling and data analysis indicated this profile as low risk probability. Responses can be Low, Medium, or High risk.



Male Previous victim Spends more than 3h on the Internet Medium empathy Medium family support In this adventure:

7. *"Playing video games in your room. Your friends joke about Pol's appearance."*

the player answers: I don't like this kind of jokes. According to you...

What is the expected risk of being a Cyber Bullying offender?

22 resp

High	5%	1 vote
Low	45%	10 votes
Medium	50%	11 votes

Figure 17. Fifth CB risk patterning question. Modelling and data analysis indicated this profile as low risk probability. This question also adds to the profiling the answer to a situation within the video game. Responses can be Low, Medium, or High risk.

Afterwards, participants were asked about their opinion regarding the **Effectiveness vs. Easiness of interventions on the available variables**. Fig. 18 shows the results obtained for this question. The variables further to the right are simultaneously easy and effective to intervene. Conversely, the variables in the lower left area are those most difficult to intervene and less effective. The score on both axes can vary from 1 (i.e., impractical/unethical) to 4 (i.e., straightforward/doable). It is also worth noting that the term "intervention" is used in a broad sense. It can include direct interventions in reality, awareness-raising campaigns focused on a specific sector of the population, awareness campaigns about the effect their actions can have on others of which they are sometimes unaware, etc. This information can be used straightforwardly to guide future interventions and educational programmes.

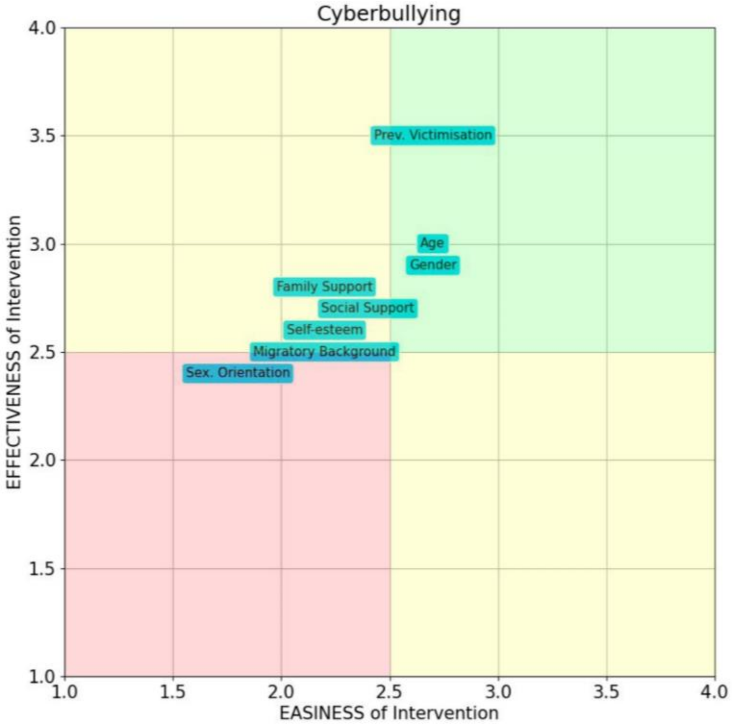


Figure 18. Workshop participants' opinion regarding Effectiveness vs. Easiness of interventions on the available variables for cyberbullying. The variables further to the top right are simultaneously easy and effective to intervene (*Age*, *Gender,* and prev. victimisation in this case). Conversely, the variables in the lower left area are those most difficult to intervene and less effective. The score on both axes can vary from 1 to 4.

3.3.1.3 Cyberbullying: Comparison of the profiling results from WP1 and WP6

Following this, the workshop proceeded with a part that focused more on discussion among the participants. They were shown tables with the points of agreement and disagreement between the results obtained (**for demographic/personal variables**) between WP6 (data analysis) and WP1 (literature review, child surveys, sentence analysis and interviews). Table 4 shows the points of agreement for the CB Offending case.

Table 4. Points of agreement in CB Offending between the analysis conducted by WP1 and WP6 for demographic/personal variables.

CB Offending WP1 & WP6 AGREEMENT
Previous CB Victimisation
High self-esteem
Gender Male
Many hours spent on the Internet
Age

About **disagreements** in CB Offending:

- In WP6 results, high **social support** appears as a prevalent risk factor, but in WP1, no such evidence was found. During the discussion, it was mentioned that this WP6 result could be compatible with the idea that perpetrators are people with a good support network. A possible explanation for an error from WP6 is that measures of low social support are sparser and noisier, so identifying them in the analyses would be more challenging. The questionnaire [4] used in the pilots may not have been able to measure this characteristic correctly, as there is an apparent shift in the distributions towards the high social value.
- The same explanation could also apply to the disagreement on **family support**. WP6 identifies high family support as a prevalent risk factor, and WP1 identifies low family support.
- WP6 results identify high self-esteem and WP1 high and low self-esteem (i.e., at the extremes). In this case, there is some agreement regarding high self-esteem. One possible explanation from WP6 is that the game has not been able to capture players with low self-esteem well (either by its design or by the setting in which the pilots take place).

Table 5 shows the points of agreement for the CB Victimisation case.

Table 5. Points of agreement in CB Victimisation between the analysis conducted by WP1 and WP6 for demographic/personal variables.

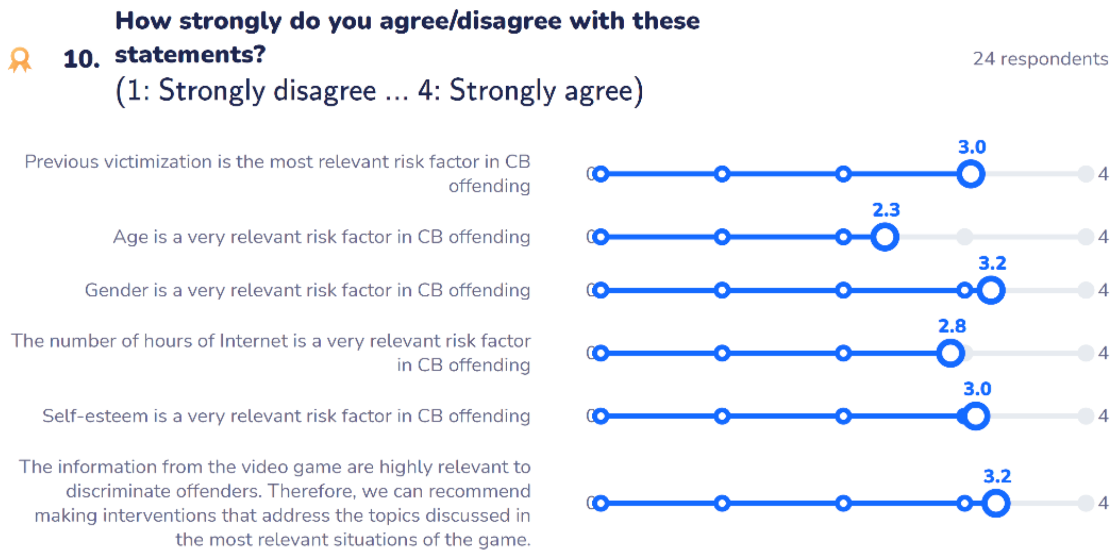
CB Victimisation WP1 & WP6 AGREEMENT
Previous CB Victimisation
Low Family Support
Age
Low self-esteem

We found **considerable disagreement** on CB victimisation (in sexual orientation, migratory background, and social support). Some possible general explanations for all these disagreements, as discussed in the workshop, are:

- The game was **specifically designed** to measure aggression and bystander, not to measure victim, so it would not work as well in this case.
- For the case of *sexual orientation*, the data has a multitude of missing values, which could be biasing the analysis. Due to the sensitivity of the question, some players might have decided to avoid answering. For the variable *migratory background*, we identified that it was not answered correctly in some countries, causing a significant bias.
- There is a huge overlap between people who have indicated Previous CB Victimization and Offending in our data. So, it is possible that we are detecting offenders indirectly in this case, but we do not have enough data to create a non-offending victim profile.

3.3.1.4 Cyberbullying: Highlights

To wrap up the cyberbullying part of the workshop, participants were given a list of potential conclusions or key points for both CB Offending (Fig. 19) and CB Victimisation (Fig. 20). They were asked to rate each statement from 1 (strongly agree) to 4 (strongly disagree). Overall, it appeared that participants generally agreed with the statements derived from the efforts of WP6 and WP1, as well as the discussions held during the workshop. The only exception was the importance of the Age variable in both CB aggression and victimisation, which received slightly lower scores.



10. How strongly do you agree/disagree with these statements?
(1: Strongly disagree ... 4: Strongly agree)

24 respondents

Previous victimization is the most relevant risk factor in CB offending — 3.0

Age is a very relevant risk factor in CB offending — 2.3

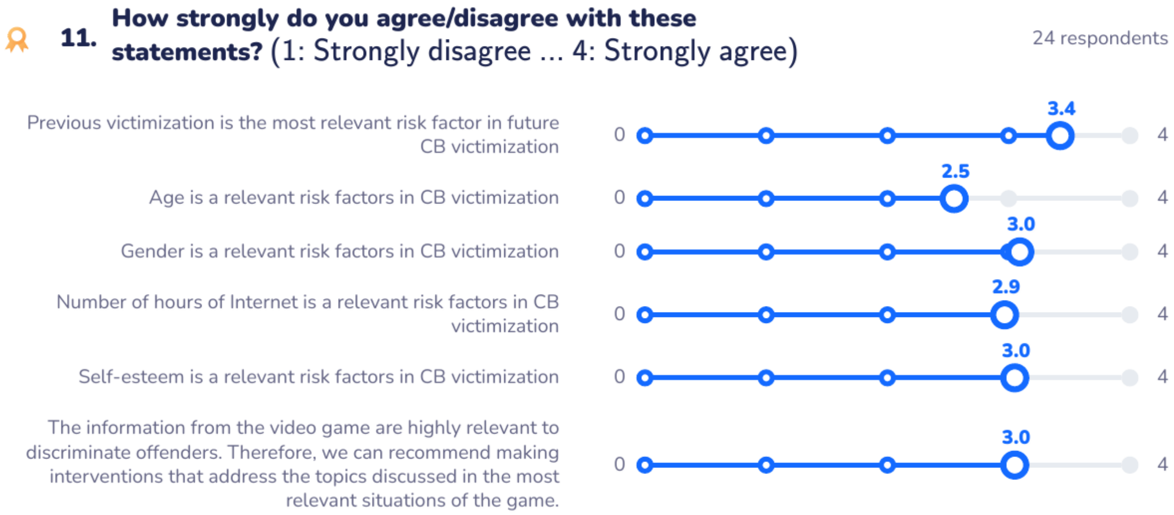
Gender is a very relevant risk factor in CB offending — 3.2

The number of hours of Internet is a very relevant risk factor in CB offending — 2.8

Self-esteem is a very relevant risk factor in CB offending — 3.0

The information from the video game are highly relevant to discriminate offenders. Therefore, we can recommend making interventions that address the topics discussed in the most relevant situations of the game. — 3.2

Figure 19. Workshop participants' average agreement with some possible highlights/conclusions obtained from the CB Offending analysis. The value ranges from 1 (i.e., strongly agree) to 4 (i.e., strongly disagree).



11. How strongly do you agree/disagree with these statements? (1: Strongly disagree ... 4: Strongly agree)

24 respondents

Previous victimization is the most relevant risk factor in future CB victimization

Age is a relevant risk factors in CB victimization

Gender is a relevant risk factors in CB victimization

Number of hours of Internet is a relevant risk factors in CB victimization

Self-esteem is a relevant risk factors in CB victimization

The information from the video game are highly relevant to discriminate offenders. Therefore, we can recommend making interventions that address the topics discussed in the most relevant situations of the game.

Figure 20. Workshop participants' average agreement with some possible highlights/conclusions obtained from the CB Offending analysis. The value ranges from 1 (i.e., strongly agree) to 4 (i.e., strongly disagree).

3.3.2 Online Grooming

In this case, unlike in the CB case, it is worth mentioning that in this cybercrime there is no *"ground truth"* or validated questionnaire with which we can verify our results. Thus, the methodology used can be interpreted as unsupervised Bayesian clustering. However, there is no guarantee that these groups correspond to children with a higher or lower risk of online grooming and that the values obtained in the experiments could be exaggerated or distorted.

3.3.2.1 Online Grooming: Summary of results from the data analysis

The results shown below summarise those obtained in deliverable D6.4, which is also openly available for anyone who wishes to consult it.

- **Research Question 1**: Which variables are most strongly related to the risk of suffering online grooming?

Based on the selected causal structure, available data, and metrics, the most relevant variables for being at risk of Online Grooming (OG) victimisation are shown in Table 6. The results of studying the influence of the variables separately (i.e., one at a time) seem to indicate that the game and the decisions that players have to make are very relevant in explaining the risk of suffering OG. No demographic/personal variables appear as relevant in this analysis.

Table 6. Most strongly related variables for the risk of suffering online grooming. In this analysis the variables are studied one at a time within the Bayesian network.

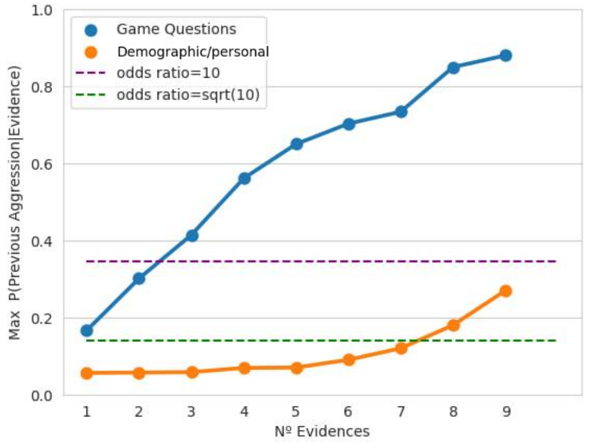
Online Grooming Victimisation Risk
Adventure 2 Question 4: Place
Adventure 2 Question 8: Friend Request
Adventure 2 Question 9: Photos
Adventure 2 Question 3: Professional type
Adventure 2 Question 5: Profile photo
Adventure 5 Question 1: Secret
Adventure 2 Question 15: Block profile

Adventure 2 Question 1: Registration Name
Adventure 5 Question 2: Biology paper
Adventure 2 Question 14: Tell parents
Adventure 2 Question 12: Ask Help
Adventure 2 Question 11: More & more
Adventure 2 Question 7: Use PC

- **Research Question 2:** What combinations of variables make it possible to construct meaningful risk profiles for suffering online grooming?

Based on the selected causal structure, available data, and metrics, the most relevant characteristics shared by the profiles at risk of suffering OG are shown in Table 7. Combining the variables in the table can create a typical OG "victim" profile. However, looking at the figure on the right of the table, we realise that **the effectiveness of the demographic/personal variables in this case is extremely low**. In other words, with the available data and the methodology used, we cannot create a truly meaningful victim profile. On the other hand, the variables from the game questions seem to show a high "profiling effectiveness".

Table 7. Summary of information obtained from the multifactorial analysis of various risk factors examined for Online Grooming victimisation risk. The colour scheme in the graph denotes variable importance, with green representing high prevalence, yellow for moderate prevalence, and red for somewhat prevalent.

Online Grooming Victimisation Risk	
Most prevalent risk factors/variables	**Evolution of "profiling effectiveness" as a function of the number of variables/evidences used**
Gender Male	
High family support	
Medium neuroticism (Big Five)	

3.3.2.2 Online Grooming: Comparison of the profiling results from WP1 and WP6

Following this, the workshop proceeded with a part of the workshop that focused more on discussion among the participants. They were shown tables with the points of agreement and disagreement between the results obtained (**for demographic/personal variables**) between WP6 (data analysis) and WP1 (literature review, child surveys, sentence analysis and interviews). Table 8 shows the points of agreement for the OG Victimisation case.

Table 8. Points of agreement in Online Grooming Victimisation between the analysis conducted by WP1 and WP6 for demographic/personal variables.

Online Grooming Victimisation Risk WP1 & WP6 AGREEMENT
Gender Male

We note that we have only found one variable where the analysis of WP1 and WP6 coincide (*Gender Male)*. According to the results of WP1, although the risk of being contacted is higher for girls, the risk of responding to such a request is higher among boys, which is what we have detected through the video game. We found **considerable disagreement** on this cybercrime. Some possible general explanations for all these disagreements are:

- As specified in the disclaimer at the beginning of Sec. 3.2.2, we have **no validated questionnaire** or "ground truth" to validate the results obtained.
- **OG is a very sensitive cybercrime to measure**, especially considering the ethical restrictions of working with minors. According to the literature, there are two types of victims: vulnerable and risky attitude. In the game, we would be measuring, indirectly, only the risky-attitude victim. So, we do not have a good direct measure of the risk of suffering from this cybercrime.
- The literature suggests that **age** may be a relevant factor. However, during the workshop, Law Enforcement Agencies (LEAs) members commented that there is a tendency towards risky attitudes at younger and younger ages. So, we might not be able to measure this difference with the ages of the participants in the RAYUELA pilots (12-16).
- Regarding measuring the **hours spent on the Internet**, the question was confusing for many children during the pilots. For example, spending 4 hours a day on social media is not the same as watching videos on *YouTube* or playing online games with your friends. This might explain why we did not observe significance in this variable. It is worth noting that we find a correlation to the question in the game where the player is put in a situation about how much time he/she has spent online. Suggesting that this might be a better measure of "self-control" with the time one spends online.

3.3.2.3 Online Grooming: Highlights

To sum up the OG segment of the workshop, participants reviewed potential conclusions about OG Victimisation risk (Fig. 21). They rated each statement from 1 (strongly agree) to 4 (strongly disagree). Overall, participants seemed to align with the statements from WP6 and WP1's work, along with the workshop discussions.

Our findings indicate that profiling based on demographic or personal traits is not very effective in OG experiments. This aligns with experts' widely shared belief that your behaviour in specific situations, rather than your physical or social traits, defines your risk tolerance. Therefore, it is safe to say that serious games provide an effective alternative for understanding how players might behave in real-life situations, given ethical considerations and appropriate game design.

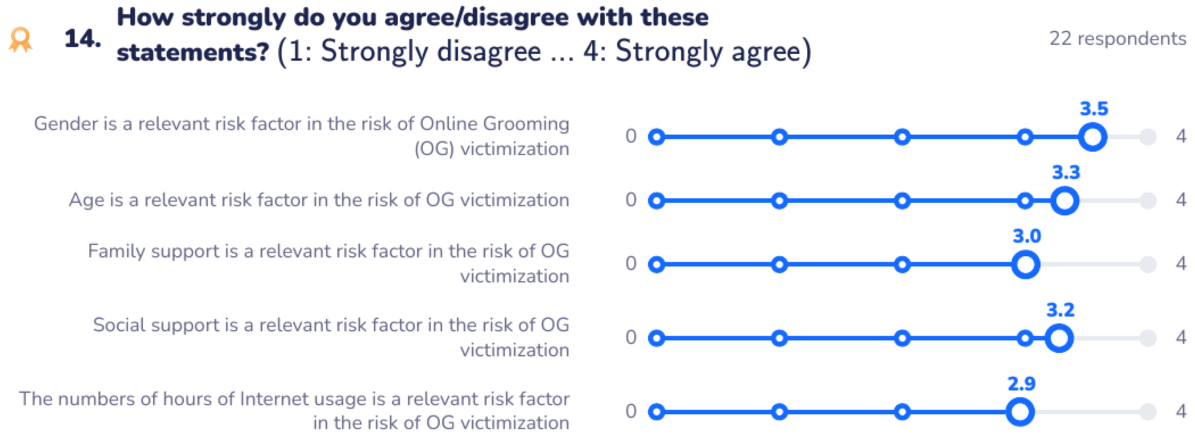


Figure 21. Workshop participants' average agreement with some possible highlights/conclusions obtained from the Online Grooming analysis. The value ranges from 1 (i.e., strongly agree) to 4 (i.e., strongly disagree).

3.3.3 Cyberthreats

As we mentioned in the case of OG, for this cybercrime, we also lack a validated questionnaire which could be used in addition to analysing game data. Thus, the methodology used can be interpreted as unsupervised Bayesian clustering. However, there is no guarantee that these groups correspond to children with a higher or lower risk of facing cyberthreats (CT), and the values obtained in the experiments could be exaggerated or distorted.

Based on the defined analytical methodology, various behaviour patterns indicating a higher risk of experiencing CT have been found. In this section, we present these results in contrast to the conclusions drawn from the literature study (WP2), which, in this case, is less extensive than that for other cybercrimes.

3.3.3.1 Cyberthreats: Summary of results from the data analysis

- **Research Question 1:** Which variables are most strongly related to the risk of suffering online grooming?

Based on the selected causal structure, available data, and metrics, the most relevant variables for being at risk of CT victimisation are shown in Table 9. The results of studying the influence of the variables separately (i.e., one at a time) seem to indicate that the game and the decisions that players have to make are very relevant in explaining the risk of suffering OG. No demographic/personal variables appear as relevant in this analysis.

Table 9. Most strongly related variables for the risk of suffering cyberthreats. This analysis studies the variables one at a time within the Bayesian network.

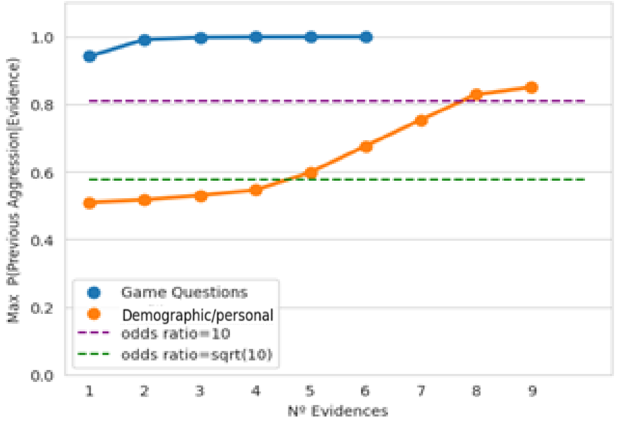
Cyberthreats Victimisation Risk
Adventure 4 Question 2: New Password
Adventure 2 Question 2: Registration Password
Adventure 4 Question 3: My Account Stolen
Adventure 4 Question 1: Phishing
Adventure 4 Question 4: Other Account Stolen

Adventure 3 Question 1: Pirated Content
"Honesty" Question

- **Research Question 2:** What combinations of variables make it possible to construct meaningful risk profiles for suffering online grooming?

Based on the selected causal structure, available data, and metrics, the most relevant characteristics shared by the profiles at risk of suffering CT are shown in Table. 10. By combining the variables in the table, we can create a typical CT "victim" profile. However, looking at the figure on the right of the table, we realise that **the effectiveness of the demographic/personal variables in this case is low**. In other words, with the available data and the methodology used, we cannot create a truly meaningful victim profile. On the other hand, the variables from the game questions seem to show a high "profiling effectiveness". This is in line with the widely held view among experts that what best defines your risk appetite is your behaviour in certain situations, not so much your physical/social/personal characteristics.

Table 10. Summary of information obtained from the multifactorial analysis of various risk factors examined for Cyberthreat victimisation. The colour scheme in the graph denotes variable importance, with green representing high prevalence, yellow for moderate prevalence, and red for somewhat prevalent.

Cyberthreats Victimisation	
Most prevalent risk factors/variables	**Evolution of "profiling effectiveness" as a function of the number of variables/evidences used**
Gender = Male	
Family Support = High	
Neuroticism = Medium	
Agreeableness = Low	
Conscientiousness = Medium	

3.3.3.2 Cyberthreats: Comparison of the profiling results from WP2 and WP6

Following this, the workshop proceeded with a part of the workshop that focused more on discussion among the participants. They were shown tables with the points of agreement and disagreement between the results obtained (**for demographic/personal variables**) between WP6 (data analysis) and WP2 (literature review). Table 11 shows the points of agreement for the CT Victimisation case.

Table 11. Points of agreement in Cyberthreats Victimisation between the analysis conducted by WP2 and WP6 for demographic/personal variables.

Cyberthreats Victimisation Risk WP2 & WP6 AGREEMENT
Gender Male
Age
Low family support
Hours spent on the Internet
Low conscientiousness (Big Five)

It is worth noting that the literature behind the analysis of WP2 is not very extensive and does not have full consistency between different sources. Nevertheless, we found a great deal of general agreement between the analysis of WP2 and WP6. We found a few **disagreements** on this cybercrime. Some possible explanations for these disagreements are:

- Most disagreements concern the **Big Five questionnaire**, for which very little evidence correlates it with the risk of CT. Thus, the results of WP2 are speculative in this regard.
- In the workshop, it was discussed that measuring personality in adolescents is complex (e.g., low self-awareness, noisy responses, high emotional lability, etc.). In addition, a version of the Big Five with fewer questions than usual was used, further increasing the noise in the answers.

3.3.3.3 Cyberthreats: Highlights

To conclude the CT part of the workshop, participants were presented with a series of possible conclusions or highlights for OG Victimisation risk (Fig. 22). Participants were asked to rate each statement from 1 (i.e., strongly agree) to 4 (i.e., strongly disagree). In general, it seems that participants agreed with the possible statements obtained from the work of WP6 and WP2, and after the discussion during the workshop. However, without much agreement either, since many of the statements are associated with demographic/personal variables, which we have explained, are not of great importance in this analysis.

As a general conclusion, we can argue that profiling through demographic/personal variables in the case of CT is ineffective in our experiments. This is in line with the widely held view among experts that what best defines your risk appetite is your behaviour in certain situations, not so much your physical/social/personal characteristics. So, we can also conclude that serious games are an effective alternative to measure how players would act in certain real-life situations (subject to ethical constraints and that the game is well designed for the desired function).

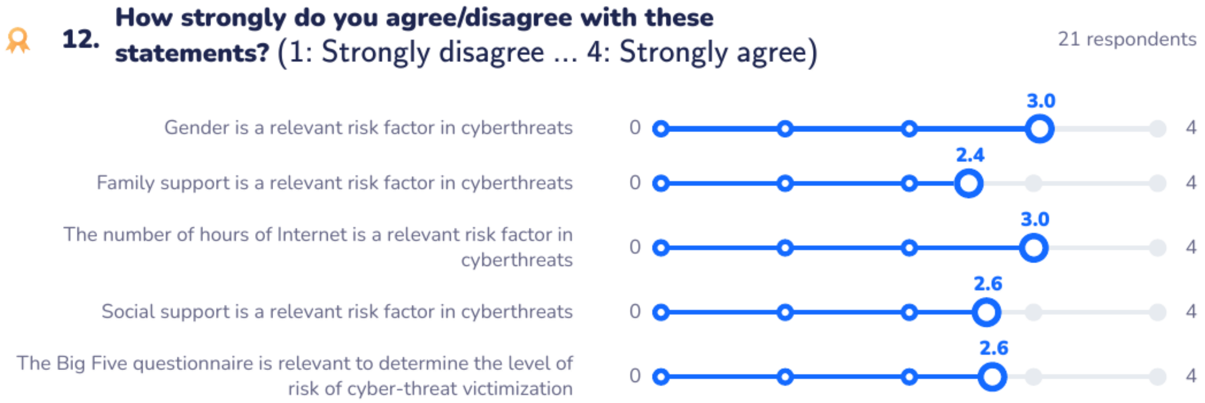


Figure 22. Workshop participants' average agreement with some possible highlights/conclusions obtained from the Cyberthreats analysis. The value ranges from 1 (i.e., strongly agree) to 4 (i.e., strongly disagree).

3.3.4 Fake News

Again, it is essential to note that in this scenario, unlike in the CB case, we lack a definitive benchmark or a verified questionnaire to cross-check our findings. Consequently, the approach employed here can be seen as unsupervised Bayesian clustering. However, it is crucial to understand that these identified groups may not accurately represent children with varying risks of spreading false information. Additionally, the values obtained in the experiments might be magnified or skewed.

The insights from the analysis offer valuable perspectives on the risk of identifying fake news among adolescents, shedding light on the factors influencing their perception and decision-making. In this section, we present these results in contrast to the conclusions obtained in the WP1. The insights from the analysis offer valuable perspectives on the risk of identifying fake news among adolescents, shedding light on the factors influencing their perception and decision-making. The expert-designed Bayesian Network demonstrated its effectiveness in capturing the intricate relationships among the available variables. This design allows us to comprehend how participants' assessments of news credibility and responses to provocative content collectively contribute to the risk of identifying fake news.

3.3.4.1 Fake News: Summary of results from the data analysis

- **Research Question 1:** Which variables are most strongly related to the risk of suffering online grooming?

Based on the selected causal structure, available data, and metrics, the most relevant variables for being at risk of Fake News (FN) victimisation are shown in Table 12. The results of studying the influence of the variables separately (i.e., one at a time) seem to indicate that the game and the decisions that players have to make are very relevant in explaining the risk of falling for FN. No demographic/personal variables appear as relevant in this analysis.

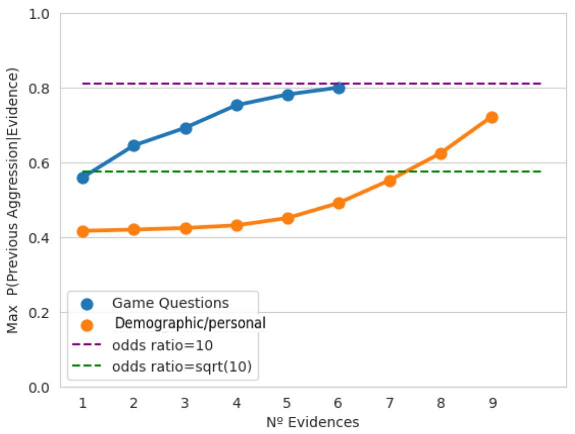
Table 12. Most strongly related variables for the risk of falling for fake news. In this analysis the variables are studied one at a time within the Bayesian network.

Fake News Victimisation Risk
Adventure 6 Question 3: Source
Adventure 6 Question 4: Information Looks Accurate
Adventure 6 Question 2: Web Page Looks Like
Adventure 6 Question 6: Regarding Charles
Adventure 6 Question 1: Migrant News Check
"Honesty" Question
Adventure 6 Question 5: Replay Post

- **Research Question 2:** What combinations of variables make it possible to construct meaningful risk profiles for suffering online grooming?

Based on the selected causal structure, available data, and metrics, the most relevant characteristics shared by the profiles at risk of falling for FN are shown in Table 13. Combining the variables in the table can create a typical FN "victim" profile. However, looking at the figure on the right of the table, we realise that **the effectiveness of the demographic/personal variables in this case is very low**. In other words, with the available data and the methodology used, we cannot create a truly meaningful victim profile. On the other hand, the variables from the game questions seem to show a high "profiling effectiveness". This is in line with the widely held view among experts that what best defines your risk appetite is your behaviour in certain situations, not so much your physical/social/personal characteristics**.**

Table 13. Summary of information obtained from the multifactorial analysis of various risk factors examined for Fake News victimisation. The colour scheme in the graph denotes variable importance, with green representing high prevalence, yellow for moderate prevalence, and red for somewhat prevalent.

Fake News Victimisation	
Most prevalent risk factors/variables	**Evolution of "profiling effectiveness" as a function of the number of variables/evidences used**
No migratory background	
Medium Neuroticism	
Gender Male	
Low Agreeableness	
Medium Conscientiousness	
High social support	
Cyberbullying Offender	

3.3.4.2 Fake News: Comparison of the profiling results from WP1 and WP6

Following this, the workshop proceeded with a part of the workshop that focused more on discussion among the participants. They were shown tables with the points of agreement and disagreement between the results obtained (**for demographic/personal variables**) between WP6 (data analysis) and WP1 (literature review).

We have not found any variable where we matched the analyses of WP1 and WP6. Some possible explanations for these **disagreements** are:

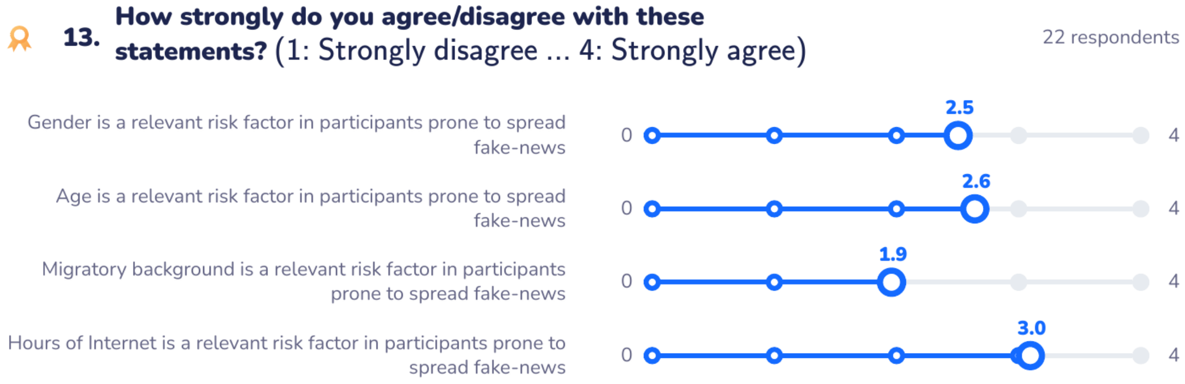
- The FN adventure was done more for educational purposes than for profiling. Therefore, it is reasonable that no meaningful profiling can be obtained.
- By design, the adventure cannot measure the risk of sharing FN, which is usually measured in the literature. Instead, it is the risk of correctly identifying a FN (e.g., verifying the information's source, cross-checking with other sources, etc.).
- Being one of the final adventures and the way the situation is set up in the game, it is possible that many participants played more exploratory without following "what they would do in reality".

3.3.4.3 Fake News: Highlights

To conclude the FN segment of the workshop, participants were given a series of potential conclusions or highlights (Fig. 23). They were instructed to rate each statement from 1 (strongly agree) to 4 (strongly disagree). Generally, it appears that the participants somewhat agree with the potential statements derived from the work of WP6 and WP2 following the workshop discussions. However, their agreement lacks firmness. Many of the statements are linked to demographic/personal factors, which, as we have explained, are not highly significant in this analysis. Moreover, the profiling analysis of this adventure is not very pertinent as it was primarily designed as an educational tool.

The critical thinking abilities of adolescents and how they assess news sources play a pivotal role in their vulnerability to fake news. Factors such as website professionalism, cautious evaluation of provocative information, and the news source's reputation influence their judgments of trustworthiness and overall susceptibility to fall for fake news. These discoveries provide valuable insights for developing educational programs and interventions to improve adolescents' media literacy and equip them with essential skills to navigate the digital landscape responsibly.

In summary, we can argue that using demographic/personal variables for profiling in the case of FN proves ineffective in our experiments. This aligns with the widely accepted expert opinion that one's risk tolerance is best defined by one's behaviour in specific situations rather than physical/social/personal traits. Consequently, we can also affirm that serious games offer a compelling alternative method to gauge how players would respond to real-life situations, provided ethical considerations are met, and the game is well-designed for its intended purpose.



13. **How strongly do you agree/disagree with these statements?** (1: Strongly disagree ... 4: Strongly agree)

22 respondents

Figure 23. Bar chart depicting the average agreement with some of the highlights/conclusions obtained from the analysis regarding Fake News. The value ranges from 1: Completely disagree to 4: Completely agree.

4. Discussion and Conclusions

In this deliverable, we collect the conclusions drawn from the data analysis carried out in WP6 and discuss them critically with the following:

(i) the results coming from WP1 and WP2,
(ii) the existing literature, and
(iii) the opinions of the other members of RAYUELA.

The main objective of the research conducted is to understand the relevance of several indicators/profiles considered on youth cybercrime. To improve the work's generalizability and the methodology's transparency, we want to emphasise the potential limitations of the work. In particular, we identify methodological limitations (e.g., algorithms or techniques used), experimental ones (e.g., biases in data collection) or interpretative ones (e.g., biases in the researchers' preconceptions in concluding).

The sample size of the data (>1000 participants and around 4000 unique adventures played) is adequate for what is usual in social science research, but it still needs to be improved. Moreover, it should be remembered that questionnaire and video game data are often noisy and heterogeneous, especially when dealing with minors (e.g., some participants may have played randomly or deliberately answered questions incorrectly).

However, the results obtained reveal exciting conclusions. The members of RAYUELA, through the workshops, have shown a general sense of consensus with the conclusions/highlights drawn. In most cases, **the variables obtained through the video game are relevant to explaining and predicting** the risk of suffering/committing cybercrimes. Therefore, the experiments indicate that the RAYUELA serious game has been correctly designed to study these cybercrimes.

Demographic variables or those obtained through psychological tests do not have significant relevance in any of the cybercrimes studied except in the case of CB offences, especially the variable indicating previous victimisation, which strongly influences the experiments. In other words, individuals who have previously been victims of CB showed a markedly higher propensity to have committed CB offences. This relationship deserves further study, although it was one of the predictions of the work developed by WP1.

Our findings indicate that attempting to profile using demographic and personal data is not fruitful in our experiments. This aligns with the prevailing expert opinion, which suggests that **one's risk appetite is better defined by one's actions** in specific scenarios rather than one's physical, social, or personal traits. Therefore, serious games are an effective alternative to measure how players would act in certain real-life situations, subject to ethical constraints and that the game is well designed for the desired function.

As indicated in the introduction to this deliverable, it should be noted that CB is the only cybercrime for which we have a "ground truth" obtained from a psychological test of cybervictimisation and cyber-aggression [4], with which to validate the results obtained. For the rest of the cybercrimes, the methodology used is close to what could intuitively be described as unsupervised Bayesian clustering. However, there is no guarantee that these clusters are smaller groups with a higher or lower risk of suffering from the cybercrime under consideration. This also means that the values obtained in the experiments might be exaggerated or distorted.

By employing **Bayesian networks** and a causal methodology (namely, Probabilistic Graphical Models), we can mitigate the biases in the data if the network structures selected are plausible. This approach also forces us to make our assumptions and hypotheses explicit, leading to discussions and critical questions about the issues we are trying to address. However, it is essential to note that **some of the methodologies** used in the experiments to "interrogate" the model once it is trained **are non-standard** since we have had to adapt them to the specific needs of our problem. This implies that some results and conclusions could be erroneous or distorted. We encourage other researchers to validate or refute our findings using different methodologies. To this end, **the data collected through RAYUELA and the serious game will be available** to everyone.

We **should be cautious with the conclusions** we can draw from the results, considering how noisy social science or video game data can be. Nevertheless, there is **room for optimism**. These results indicate that **the serious game has been designed correctly** and that, through the responses gathered, we can statistically discern between victims/non-victims of the considered cybercrimes. Further research and more data are necessary to validate and generalise these findings. By doing so, we can create more effective preventive measures to mitigate cybercrime in minors and ensure the safety of society.

References

[1] Rammstedt, B., & John, O. P. (2007). Measuring personality in one minute or less: A 10-item short version of the Big Five Inventory in English and German. Journal of research in Personality, 41(1), 203-212.

[2] Rosenberg, M. (1965). Rosenberg self-esteem scale (RSE). Acceptance and commitment therapy. Measures package, 61(52), 18.

[3] Zimet, G. D., Dahlem, N. W., Zimet, S. G., & Farley, G. K. (1988). The multidimensional scale of perceived social support. Journal of personality assessment, 52(1), 30-41.

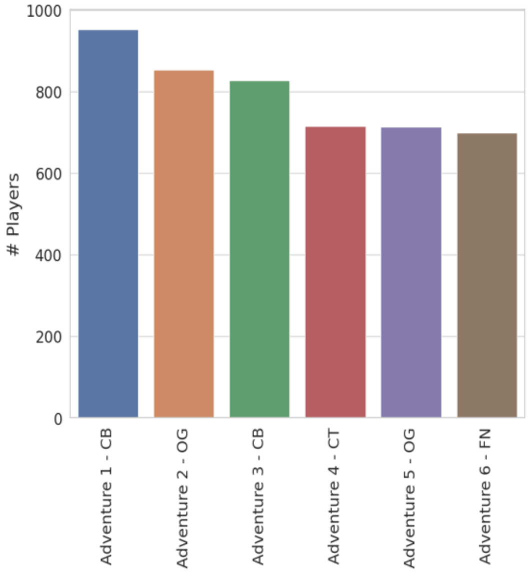
[4] Brighi, A., Ortega, R., Pyzalski, J., Scheithauer, H., Smith, P. K., Tsormpatzoudis, H., Tsorbatzoudis, H., et al. (2012). European Cyberbullying Intervention Project Questionnaire (ECIPQ) [Database record]. APA PsycTests.

Annex 1: Exploratory Data Analysis

Below is a set of descriptive statistics on the data collected up to the third phase of the RAYUELA pilots. Gameplay data has not been considered in this exploratory analysis. We only considered demographic data and the psychological/sociological questionnaires that the students had to fill in before and after playing the game.

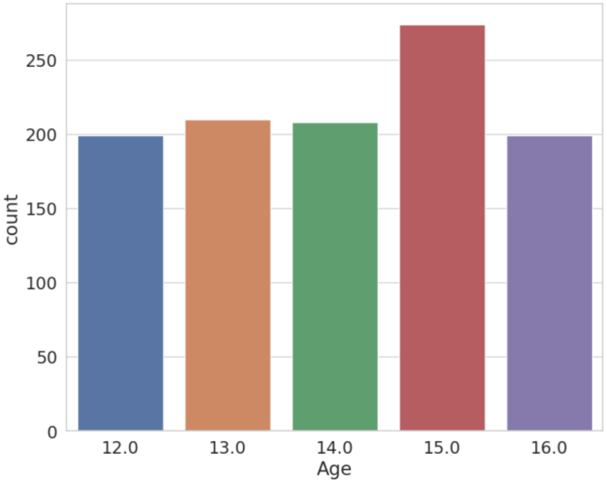
Number of players

Total Players in the registry: 1147		
Adventure	# Players	Percentage
Adventure 1 - CB	953	83.1%
Adventure 2 - OG	853	74.4%
Adventure 3 - CB	828	72.2%
Adventure 4 - CT	716	62.4%
Adventure 5 - OG	714	62.3%
Adventure 6 - FN	699	60.9%



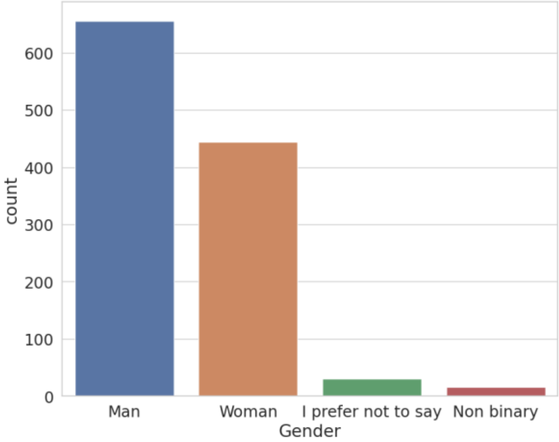
Age

Age	# Players	Percentage
12	199	17.35%
13	210	18.3%
14	208	18.1%
15	274	23.9%
16	199	17.35%



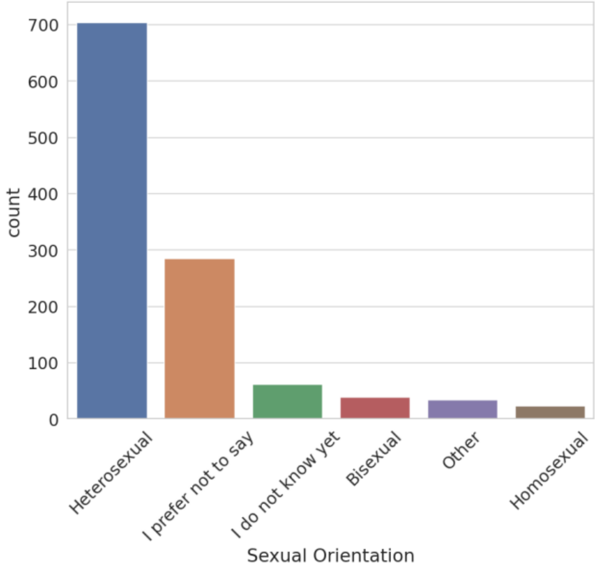
Gender

Gender	# Players	Percentage
Man	656	57.2%
Woman	444	38.7%
“I prefer not to say”	31	2.7%
Non-Binary	16	1.4%



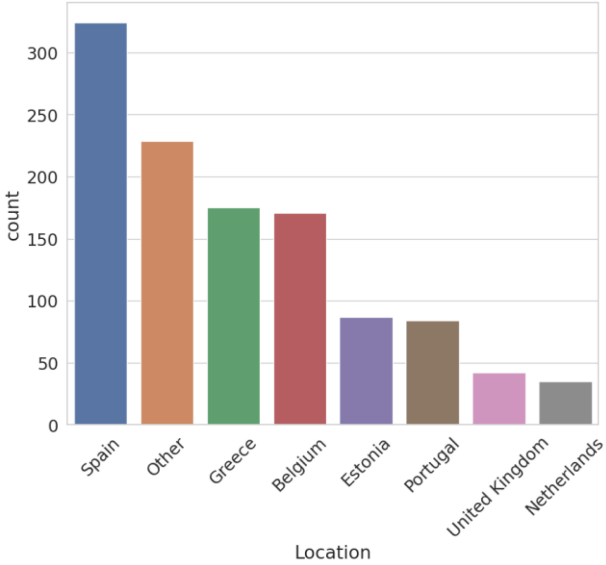
Sexual Orientation

Sexual Orientation	# Players	Percentage
Heterosexual	704	61.4%
“I prefer not to say” / Not asked	285	24.8%
“I don’t know yet”	62	5.4%
Bisexual	39	3.4%
Other	34	3%
Homosexual	23	2%



Country

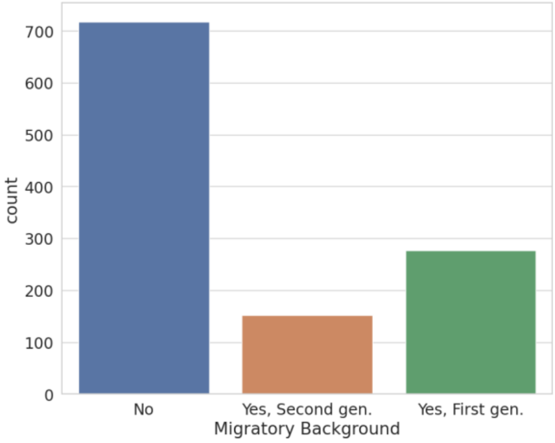
Sexual Orientation	# Players	Percentage
Spain	324	28.25%
Other	229	20%
Greece	175	15.25%
Belgium	171	14.9%
Estonia	87	7.6%
Portugal	84	7.3%
United Kingdom	42	3.7%



Netherlands	35	3.05%

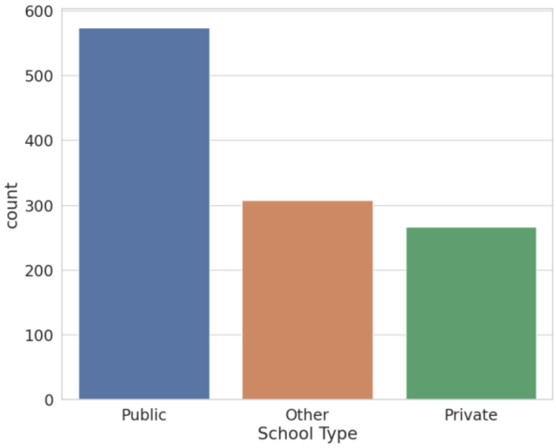
Migratory Background

Migratory Background	# Players	Percentage
No	718	62.6%
Yes, First Gen.	277	24.15%
Yes, Second Gen.	152	13.25%



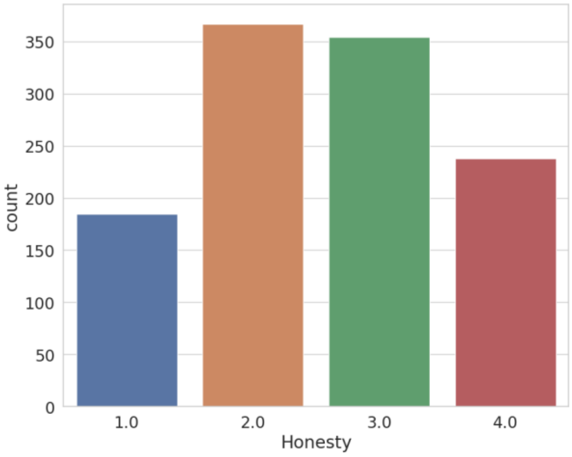
School Type

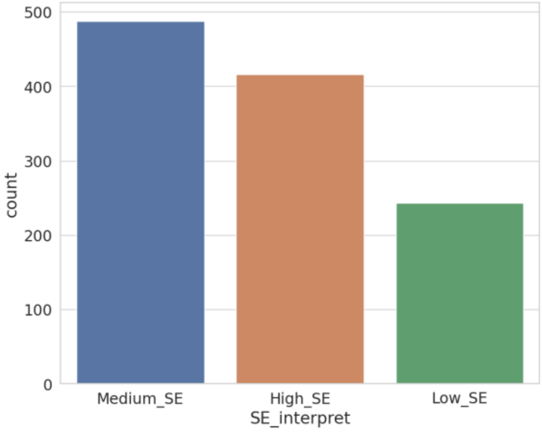
Migratory Background	# Players	Percentage
Public	574	50%
Other	307	26.8%
Private	266	23.2%



"Have you played like you would behave in real life?"

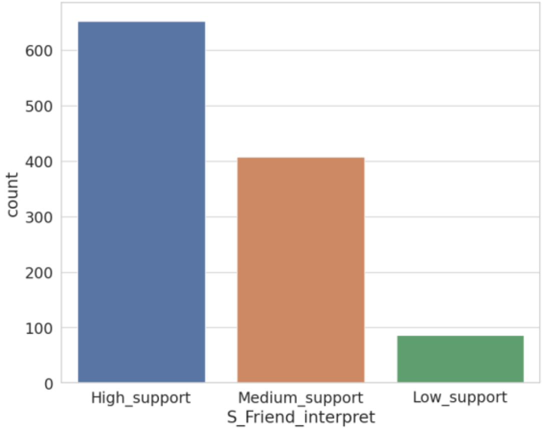
Have you played like you would behave in real life?	# Players	Percentage
1 – very different	185	16.2%
2 – different	367	32.1%
3 – similar	354	30.9%
4 – very similar	238	20.8%





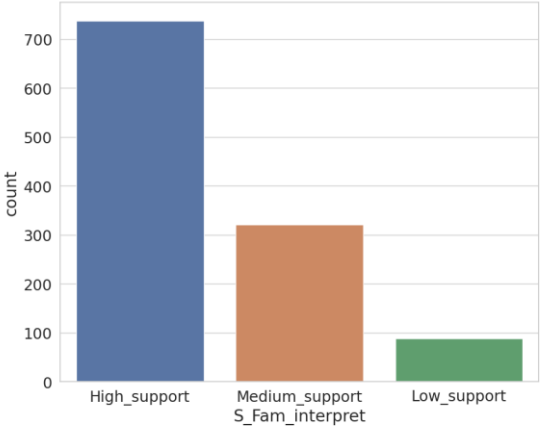
Self-Esteem

Self-Esteem	# Players	Percentage
Medium	488	42.5%
High	416	36.3%
Low	243	21.2%



Social Support

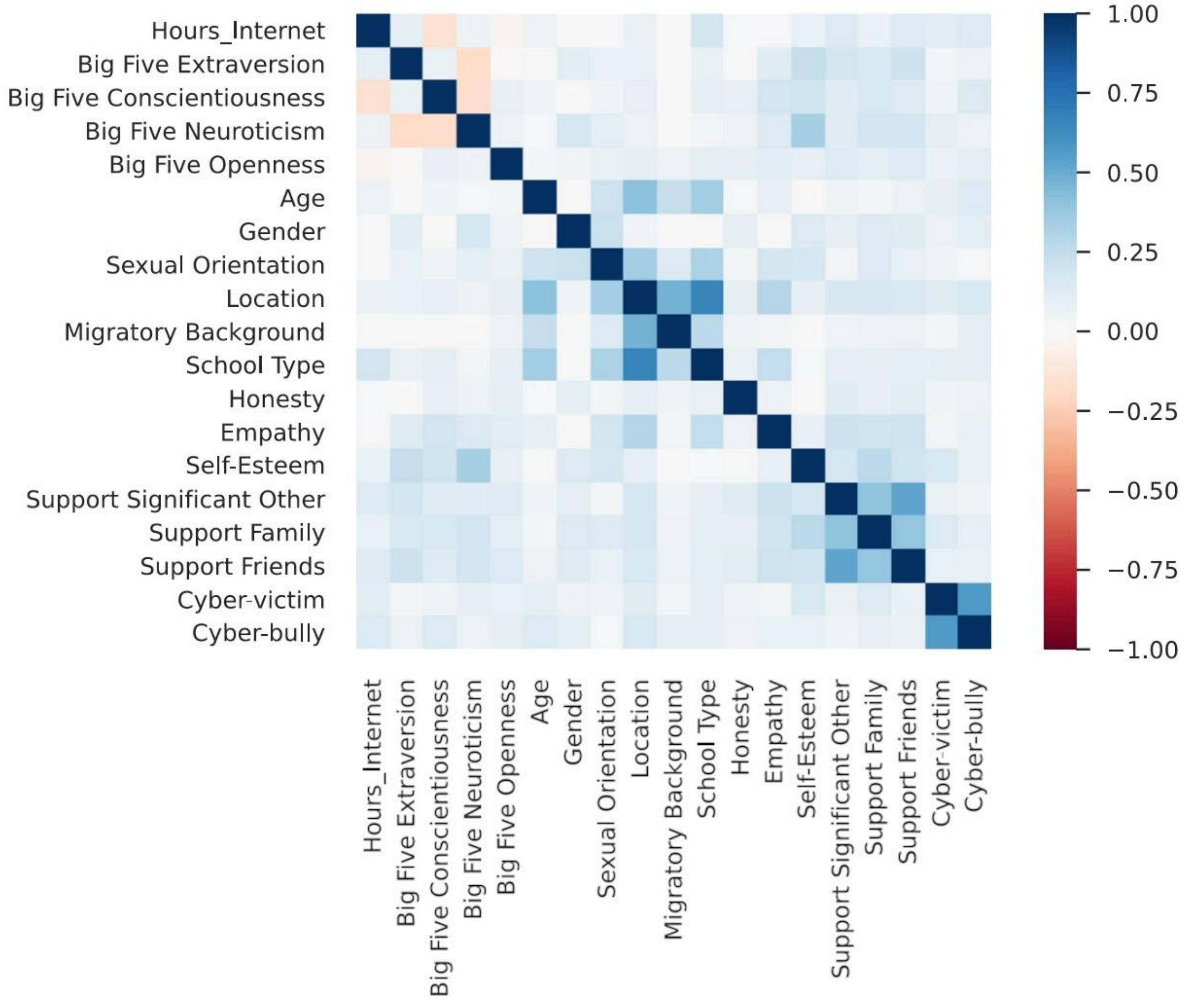
Social Support (Friends)	# Players	Percentage
High	653	56.9%
Medium	408	35.6%
Low	86	7.5%



Family Support

Social Support (Friends)	# Players	Percentage
High	738	64.3%
Medium	321	28%
Low	88	7.7%

Correlations between variables



Significant correlations:

- “Location” is highly correlated with “School type”
- “Cyber-victim” is highly correlated with “Cyber-bully”
- “Support Friends” is highly correlated with “Support Significant Other”