1 *Article*

2 Adolescents' Exposure to Online Risks: Gender
3 Disparities and Vulnerabilities Related to Online
4 Behaviors

5 **Elena Savoia, [1,2,*] Nigel Harriman, [1] Max Su, [1,2] Tyler Cote, [3] and Neil Shortland [4]**

6 [1] Emergency Preparedness, Research, Evaluation, and Practice Program, Division of Policy Translation and Leadership
7   Development, Harvard T.H. Chan School of Public Health, Boston, MA 02115, USA
8 [2] Department of Biostatistics, Harvard T.H. Chan School of Public Health, Boston, MA 02115, USA
9 [3] Operation 250, Lowell, MA 01854, USA
10 [4] Center for Terrorism and Security Studies, University of Massachusetts Lowell, Lowell, MA 01854, USA
11
12 [*] Author to whom correspondence should be addressed: preparedness@hsph.harvard.edu

13 **Abstract**: In the last decade, readily available electronic devices have created unprecedented opportunities
14 for teens to access a wide variety of information and media – both positive and negative – on the internet.
15 Despite the increasing number of initiatives taking place worldwide intended to assess and mitigate the
16 online risks encountered by children and adolescents, there is still a need for a better understanding of how
17 adolescents use the internet and what consequences they may face. We conducted a cross sectional online
18 survey of a convenience sample of 733 8th and 9th grade students in Utah. The survey contained eight
19 questions regarding students' exposure to three types of online risk scenarios: content risk, contact risk, and
20 commercial risk. Independent variables included students' online behaviors, use of social media and private
21 messaging apps, and adult supervision of online activities. Logistic and negative binomial regression models
22 indicated that female gender, social media use, and chatting with strangers were associated with exposure
23 to multiple unsafe online scenarios. Our results provide critical information to practitioners involved in the
24 development of educational initiatives by building a profile of potentially risky online behaviors and allow
25 them to tailor their initiatives to meet the needs of more vulnerable populations.

26 **Keywords:** Online Safety; Online Risk; Online Behaviors; Gender Disparities; High School Students; Social
27 Media

28

29 1. Introduction

30   In the last decade, readily available electronic devices such as computers, tablets, and smartphones have
31 created unprecedented opportunities for teens to communicate, connect, share, and access a wide variety of
32 information and media on the internet [1, 2]. A 2018 survey reported that nearly all U.S. teens (95%) have access
33 to a smartphone and 45% are "almost constantly" on the internet. Information and communication
34 technologies are playing a key role in fulfilling adolescents' emotional and communication needs [3], and the
35 instrumental and social functions of such technologies are having a critical impact on teens' interactions with
36 parents, peers, and social groups [4]. Digital communication can both strengthen and strain teen-parent
37 relationships. Mobile phones provide a permanent channel of communication between parents and teenagers,
38 thereby offering opportunities to intensify parental supervision and control [5-7]. However, the same technology
39 is also facilitating the process by which teens connect with peers and with individuals in an independent
40 manner free from parental surveillance.
41   Despite the increasing number of initiatives taking place worldwide intended to assess and mitigate the
42 online risks encountered by children and adolescents, there is still a need for a better understanding of how
43 adolescents use the internet and what consequences they may face [2, 8]. The current understanding of how
44 young people experience communication technologies, how they perceive opportunities for risks over the
45 internet, and how social media shapes their interactions is developing but admittedly incomplete.

1.1. Adolescents' exposure to risky online situations

Risky online situations that adolescents may encounter have been classified into three groups: 1) "content risk", referring to user exposure to possible harmful contents, 2) "contact risk", referring to user activities and communications with known or unknown individuals that carry a potential threat and 3) "commercial risk", referring to situations in which commercial organizations attempt to exploit the user [9]. As a highly interactive platform, social media facilitates more opportunities for adolescents to engage in risky behaviors. When online, adolescents devote much of their free time to social media, using social media platforms for an average of nearly 3 hours each day [10, 11]. Understanding the relationship between social media use and risky behaviors during adolescence is crucial given the increased propensity for risk-taking that is unique to this developmental period [12, 13]. Several studies and reviews of existing literature provide evidence illustrating how the use of social media impacts – both positively and negatively – essential aspects of adolescents' mental health and psychosocial development, including self-esteem, social connectedness, peer communications [14-16], and how social media may pose risks of cyber victimization and lurking [17-20]. A recent meta-analysis showed that higher levels of social media use among adolescents are associated with more frequent engagement in the use of drugs, risky sexual acts, and violent behaviors with a small-to-medium magnitude of effect [21]. Specific examples of negative situations experienced by adolescents in the online space include: being contacted by a stranger, having either seen or received media that made them feel uncomfortable, having felt under pressure to send photos or other information about themselves, and having accidentally spent money online that they did not mean to spend [22, 23]. While all these situations have been widely documented, little is known about which online behaviors and which social media and communication platforms are most likely associated with negative and risky situations.

1.2. Current study

This study, seeking to explore predictors of adolescents' exposure to online risks, is guided by the following exploratory research questions: What demographic characteristics are associated with exposure to online risks? How is the use of social media and private messaging apps associated with exposure to online risks? Is adult supervision of adolescents' online activities a protective factor towards exposure to online risks?

2. Materials and Methods

2.1. Study Design

We implemented a cross-sectional study design by the use of an online survey, in four schools in Utah (USA), gathering data from students in 8th and 9th grade. The survey was implemented in April 2018 and parents were provided with opt-out forms one month before the start of data collection. Data were collected using the online survey platform Qualtrics. The survey link was distributed to the students by their teachers during class time – the survey took roughly twenty minutes to complete. The study protocol and instruments were approved by the Harvard T.H. Chan School of Public Health Institutional Review Board (IRB16-1757) as well as by the ethical committee of the school district where the study was implemented. The Helsinki ethics protocol was followed throughout the course of this study [24]. A copy of the survey instrument can be found in Appendix A.

2.2. Dependent Variables

To measure exposure to online risks, respondents were asked if they had ever experienced any of the following 8 situations: 1) Bullying or harassment by friends or acquaintances, 2) Getting involved in unwanted conversations in a chat room, social networking site, or on email, 3) Coming across sexual images or content, 4) Coming across images of violence, 5) Someone trying to sell me drugs or alcohol, 6) Someone using my photos in an inappropriate way, 7) A stranger trying to meet with me, 8) Coming across hate groups trying to convince me of their views. Some of these situations can be classified as content risk (q3 and q4), some others as contact risk (q1, q2, q6, q7, and q8), and one as commercial risk (q7). This dependent variable was studied in two ways: 1) using each of the eight situations as an outcome, 2) using the number of situations a student encountered as the outcome. We also created a dichotomous variable where 0 = no situations encountered and

93 1 = any number of situations encountered, to further study and confirm the relationship found in the previous
94 analysis.

2.3. Independent variables

96 Independent variables included demographics (race, gender, and age), in addition to academic
97 performance, use of social media and private messaging apps, and adult supervision. Academic performance
98 was measured by frequency of grade type (A, A-/B+, B, B-/C+, and C or lower). Use of social media was
99 measured by asking respondents which social media tool (i.e. YouTube, Snapchat, Instagram, etc.) they use
100 and how frequently they use it, with answer options ranging from 1 (never) to 6 (all the time). A summative
101 social media use score was created by dichotomizing the use of each social media platform used by the
102 respondent into a dichotomous variable (1=daily use or higher, 0=less than daily use) and summing the
103 attributed value across all fifteen platforms. Respondents were also asked if they used any private messaging
104 apps, including Kik, Telegram, or WhatsApp, and if they used such apps more frequently than texting.
105 To further understand social media use and online behaviors, respondents were asked how many of their
106 social media followers they knew in person, and how frequently they chatted with strangers while on social
107 media. A similar question was asked about their friends' behaviors, under the assumption that they would be
108 less likely to misreport friends' habits compared to their own. Additional questions focused on behaviors such
109 as sharing personal information while on social media or chatting with strangers while playing online video
110 games. Finally, respondents were asked about adult supervision, in particular, if their parents supervised their
111 online activities, if anyone had ever talked to them about online safety, and if they had a trusted adult to ask
112 for help in case they encountered an uncomfortable situation online.

2.4. Statistical Analyses

114 Simple and multiple logistic regression models were used to study how the independent variables were
115 associated with the occurrence of each of the 8 situations identified as online risks. A Box-Tidwell procedure
116 was used to confirm that aggregate social media use had a linear relationship to the log odds of each of the 8
117 dependent variables. Independent variables were included in the multiple models when a statistically
118 significant association was found in the simple models (p-value < 0.05). Gender, race, and age were included
119 regardless of their significance in the final models because of theoretical relevance. Hosmer-Lemeshow tests
120 were used to assess the goodness of fit of the multiple logistic regression models [25]. For the second dependent
121 variable related to the number of situations the respondents experienced, a negative binomial regression
122 model was fitted to the counts of the unsafe online situations. Pearson's chi-squared test of dispersion was
123 used to assess the goodness of fit for the negative binomial model. We compared the fit of the negative
124 binomial regression model to Poisson, zero-inflated Poisson, and zero-inflated negative binomial models
125 using the Bayesian information criterion (BIC) to determine whether any of those distributional forms was
126 more appropriate for this sample. For both logistic and negative binomial regressions, model misspecification
127 was assessed using Pregibon's goodness-of-link test [26]. For models where the goodness-of-fit test failed, we
128 tried to modify the set of covariates in the model to achieve a better fit as feasible. Data analysis was performed
129 using Stata Statistical Software: Release 15.1. College Station, TX: StataCorp LLC.

3. Results

131 The response rate to the survey was 86% (767/897). This paper is based on the analysis of 733 students for
132 which we had complete data on the dependent variable. Below we provide the sample characteristics based
133 on demographics and academic performance, and descriptive statistics for social media and private messaging
134 use, social media behaviors, adult supervision of online activities, and unsafe online situations experienced by
135 the respondents. Following, we present the results of the multiple logistic regression models for the two
136 dependent variables (each of the 8 unsafe situations and exposure to at least one situation). Finally, we present
137 the results of the negative binomial regression model for the count of unsafe online scenarios.
138

Table 1 about here

140 *3.1. Sample characteristics and descriptive statistics*

141 Detailed descriptive statistics for the sample can be found in Table 1. We gathered data from a
142 convenience sample of 733 individuals, the majority of whom were male (51%). Major categories of race were:
143 White (39%), Hispanic (21%), and mixed-race (10%). The sample included students aged 14 (41%), 15 (58%),
144 and 16 (1%). Regarding academic performance, the majority of students reported receiving grades ranging
145 between A- and B+ (41%).

146 3.1.1. Social media use and online behaviors

147 Detailed descriptive statistics for the sample can be found in Table 1. Most participants reported using
148 more than one social media platform on a daily basis (85%), and 26% reported using three social media
149 platforms daily. Only 14% of students had met all of their social media followers in person, the majority of
150 them (62%) reported that they had chatted with strangers on social media, and 62% reported they believe their
151 friends did so as well. Thirty-one percent of respondents indicated they had shared personal information, such
152 as their school or town name when posting on social media. Fourteen percent of students used private
153 messaging apps, and among these students most frequently used WhatsApp (67%), followed by Kik (37%),
154 and Telegram (9%). The majority of participants played video games online (78%), and among those who did,
155 59% reported that they chatted with someone they did not know while gaming. Regarding parents'
156 supervision, the majority (52%) of students reported that their parents did not install parental controls on their
157 computer or other devices. The majority (69%) of students reported they had a trusted adult they could ask
158 for help if they experienced an online situation that made them feel uncomfortable. Finally, 95% of respondents
159 reported that someone had spoken to them about online safety.

160 3.1.2. Unsafe online situations

161

162 **Table 2 about here**

163

164 Table 2 presents the detailed summary measures of each unsafe online event and the count of unsafe
165 online events. The most frequently reported situation was coming across images of violence (32%), followed
166 by coming across sexual images or content (31%), and bullying by friends or acquaintances (29%). Twenty-six
167 percent of respondents indicated they had been in an unwanted conversation in a chat room, social networking
168 site, or on email. Only six percent of students reported someone using their photos in an inappropriate way.
169 Fourteen percent of students indicated that someone had tried to sell them drugs or alcohol online. Seventeen
170 percent of students reported experiencing a situation in which they encountered a stranger online who wanted
171 to meet with them, and 14% of students reported coming across hate groups trying to convince them of their
172 views. Overall, the distribution of the counts of unsafe online situations was skewed to the right with subjects
173 reporting a mean of 1.7 events (SD = 2) and a median of 1 (range=0-8). The majority of the subjects (40%) did
174 not report experiencing any of the unsafe online situations listed in the survey.

175 *3.2. Logistic Regression Models*

176

177 **Table 3 about here**

178

179 Table 3 presents the results for the multiple logistic regression models for each unsafe online event.
180 Simple logistic regression models were performed for each of the dichotomous dependent variables. Box-
181 Tidwell test results confirmed that our summative measure of overall social media use was linear to the log

182   odds of each outcome (p>0.05). Detailed results for the simple models can be found in Appendix B. The overall
183   LR chi-square test statistics for the 8 multiple models exploring the association between the independent
184   variables and each of the 8 online situations were significant (χ^2, p<0.01). Hosmer-Lemeshow Goodness of Fit
185   test results confirmed that all models were a good fit for the data. Pregibon's link test did not indicate any
186   model misspecification ($\hat{y}$ p<0.05, $\hat{y}^2$ p>0.05), with the exception of the dependent variable someone trying to
187   sell the respondent drugs. Below we provide a summary of key results for each situation and the counts of
188   situations experienced.

189   3.2.1. Online content risk

190   Two of the eight unsafe online situations we investigated referred to exposure to harmful contents
191   (content risk): 1) coming across sexual images or content and 2) coming across images of violence. Gender was
192   predictive of exposure to both of these risks. Compared to males, females had twice the odds of coming across
193   sexual images or content (OR=2, 95% CI 1.4-2.9) and 1.9 times the odds of coming across images of violence
194   (OR=1.9, 95% CI 1.3-2.7). Chatting with a stranger while online was also a predictor for both of these content
195   risks. Respondents who reported this behavior had 70% increased odds of coming across sexual content
196   (OR=1.7, 95% CI 1.1-2.5) and images of violence (OR=1.7, 95% CI 1.1-2.6). Race was the only other variable that
197   predictive of exposure to sexual images; white students had 50% increased odds of coming across the same
198   content compared to non-white students (OR=1.5, 95% CI 1-2.1).
199   Regarding the other content risk, age, academic performance, and social media use predicted coming
200   across images of violence. Students aged 15 and older had 70% increased odds of coming across images of
201   violence (OR=1.7, 95% CI 1.2-2.4). Students who received grades of B or lower had 60% increased odds of
202   coming across images of violence, compared to those who had grades of B+ or higher (OR=1.7, 95% CI 1.2-2.4).
203   Finally, the odds of coming across images of violence increased by 24% with each additional social media
204   platform used by the respondent on a daily basis (OR=1.24, 95% CI 1.08-1.43).

205   3.2.2. Online contact risk

206   Five out of the eight online unsafe situations we investigated referred to interactions with someone online
207   that lead to an unsafe outcome (contact risk situations): 1) bullying and harassment by friends or
208   acquaintances, 2) getting involved in unwanted conversations, 3) someone using the respondent's photos in
209   an inappropriate way, 4) a stranger trying to meet with them, and 5) coming across hate groups trying to
210   convince the respondent of their views. Chatting with a stranger while online was a predictor for all of the five
211   contact risks. Respondents who reported this behavior had increased odds of exposure to the five contact risk
212   situations. These increased odds ranged from 80% for having experienced unwanted conversations (OR=1.8,
213   95% C.I. 1.2-2.9), and for coming across hate groups (OR=1.8, 95% C.I. 1-3.2) to 350% for having had someone
214   using the respondent's photos in an inappropriate way (OR=4.5, 95% C.I. 1.5-13.3). Gender was associated
215   with two of the five situations, with females having 70% increased odds of experiencing bullying or
216   harassment while online (OR-1.7, 95% 1.2-2.4) and 100% increased odds of engaging in unwanted
217   conversations (OR=2, 95% C.I 1.4-2.9). The number of social media platforms used by the respondent was
218   associated with having experienced bullying or harassment the online space, the odds of experiencing bullying
219   or harassment increased by 14% with each additional social media platform used by the respondent on a daily
220   basis (OR=1.14, 95% CI 1.01-1.28). Finally, the use of the private messaging app Kik was associated with 210%
221   increased odds of having someone used the respondent's photos in an inappropriate way (OR=3.1, 95% CI 1.1-
222   9.2).

223   3.2.3. Commercial risk

224   One of the eight situations we investigated refers to the commercial risk: *someone trying to sell the*
225   *respondent drugs or alcohol.* None of the independent variables resulted to be a significant predictor of this
226   experience.

227   *3.3. Negative binomial model for number of types of unsafe online situations*

228

229   **Table 4 about here**

230

231   Table 4 displays the results of the negative binomial model for the number of types of unsafe online
232   situations. The final negative binomial regression model had a Pearson's chi-square of 739.8 (df=693, *p*=0.11)
233   indicating the model accounted for any overdispersion in the data well. Also, Pregibon's link test did not
234   indicate any model misspecification ($\hat{y}$ *p*<0.01, $\hat{y}^2$ *p*=0.36). The model included the following predictors: age,
235   gender, race, social media use, chatting with strangers on social media, the respondent's friends' habit of
236   chatting with strangers on social media, parents' supervision of online activities, if anyone had ever talked to
237   them about online safety and if they have a trusted adult to ask for help with unsafe online situations. The
238   negative binomial model had a BIC of 2445.6 which was favorable when compared to the BIC from a Poisson
239   (BIC=2646.4), zero-inflated Poisson (BIC=2487.5), and zero-inflated negative binomial (BIC=2460.7) models
240   with similar predictors.
241   In the negative binomial model, females had 1.5 times the incidence rate of experiencing different types
242   of unsafe online situations compared to males (IRR=1.5, 95% CI 1.3-1.9). Social media usage was also a
243   significant predictor with each additional social media platform used on a daily basis associated with an 8%
244   increase in the incidence rate of experiencing different types of unsafe online scenarios (IRR=1.08, 95% CI 1.01-
245   1.15). Students who chatted with strangers on social media had 1.6 times the incidence of experiencing
246   different types of unsafe online scenarios than those who did not (IRR=1.6, 95% CI 1.3, 2.0). Lastly, the
247   frequency in which the respondent's friends chatted with strangers on social media was a significant predictor
248   (χ^2(df)=14.6(3), *p*=0.002). This was driven by significantly higher incidence rates of experiencing different types
249   of unsafe online scenarios for students whose friends "sometimes" or "often" chatted with strangers on social
250   media compared to students who did not know if their friends chatted with strangers with an IRR of 1.5 (95%
251   CI 1.2-1.9) and 1.6 (95% CI 1.2-2.1), respectively.
252   Similar results were obtained in logistic regression models for our dichotomized variable for exposure to
253   any unsafe online situation (see Table 3). Results from these analyses similarly indicated that gender (OR=2.0,
254   95% CI 1.4-2.7), social media use (OR=1.16, 95% CI 1.01-1.32), chatting with strangers online (OR=2.3, 95% CI
255   1.6-3.4), and friends chatting with strangers online (χ^2(df)=11.8(3), *p*=0.008) was significantly associated with
256   exposure to unsafe events online.

257   **4. Discussion and Conclusions**

258   The goal of this study was to explore the predictors of adolescents' exposure to online risks, with a specific
259   focus on the demographic characteristics that are associated with exposure to online risks, which social media
260   and private messaging apps are associated with exposure to online risks and how/if adult supervision of
261   adolescents' online activities a protective factor towards exposure to online risks. Overall, our findings support
262   the important role that certain individual characteristics (namely gender and age) play as a risk factor for
263   exposure to risk online. Findings from our study show how some children appear to be particularly "at risk"
264   when online; being more likely to encounter risky or uncomfortable situations when on the internet. In
265   particular, our study points to the existence of gender disparities in the exposure to online risks, with girls
266   being more likely than boys to encounter online risks. The role of gender is inevitably complex and
267   multifaceted, but our data show that such disparities seem to be independent of other online behaviors such
268   as the amount of time spent online or type of social media platform being used, suggesting how social factors
269   might play a role including the potential for girls being the target of harassment, violence and sexual abuse.
270   In addition to this, the type and duration of social media activity did increase some forms of online risk.
271   Finally, there was little evidence of parental supervision as a protective factor.
272   Overall, the theoretical importance of this paper is to focus on the role of the individual as a precursor to
273   being exposed to risk online. At the global level, we are increasingly focused on the harm caused by material
274   accessed on the Internet. Policymakers, academics, politicians, mental health, and medical professionals seek
275   to understand the processes online that govern the relationship between internet use and risk. While the
276   majority of research has explored the effect of the Internet on the individual – such as the induced feeling of
277   anonymity – [27], increasingly, research is exploring the role of the *user* on the likelihood that they engage in or

278   are exposed to risky material online [28, 29]. This research reinforces this interactionist perspective of online risk
279   in that risk is the result of an interaction between the nature of "being online" with the pre-existing
280   characteristics and personality of the user.

281   In addition to this, these findings also support the idea of differential risk in that while certain
282   characteristics have a domain-general effect on risk, increasing the chance of engaging in risky behavior across
283   a range of behaviors, others increase risk in a refined (or domain-specific) manner. For example, while being
284   female increases the risks of being exposed to bullying, violent and sexual images, unwanted conversations,
285   and at least one risky event, using the private messaging app *Kik* only increased the chances of an individual's
286   photos being used inappropriately. From a theoretical perspective, what this encourages is the need to better
287   differentiate the relationship between characteristics, and or social media use and the specific type of risk that
288   someone is engaged in. From a behaviorist standpoint, extant literature has shown that broad patterns of
289   behavior reflect different fundamental incentives to human action [30, 31]. What this means then is that when it
290   comes to online crime, we need to unpack the underlying incentives that guide both engagement with certain
291   forms of online behavior and the certain risks that may stem from this.

292   These findings thus have important implications for the prevention of online harm. The question of
293   coping is important as part of a wider shift from a technologically determinist discourse (of what the internet
294   "does" to children) in favor of recognizing the importance of building children's digital resilience.
295   Encountering risk could in itself also represent an opportunity – to become more resilient, more digitally
296   literate, and less vulnerable to online risks [32]. Sustained efforts are required from a variety of stakeholders,
297   including families, schools, and technology industries in order to effectively safeguard children online in a
298   dynamic technological and social media landscape. Thus, it is important to identify vulnerabilities that expose
299   some children to higher risks so as to focus prevention efforts on those who will most benefit from them.

300   Educational initiatives are then tasked with developing programming for youth to address these
301   underlying influences associated with online threats. These initiatives should be designed to develop the
302   critical capability in children to recognize and mitigate the risks encountered by also learning how to self-
303   regulate their behavior and seek help when needed. Children learning how to identify and interpret the impact
304   and potential repercussions of various online behaviors can be influential in the overall safeguarding of youth
305   online; among these potential risks, our study highlighted those posed by chatting with strangers, excessive
306   use of social media platforms, and private messaging apps, and the development of both the technical and
307   emotional competence to self-regulate online behaviors and deal with uncomfortable situations.

308   Initiatives that build knowledge and awareness about online risks are necessary, as is better awareness
309   of existing disparities in exposure to risks. What this research can lend to the building and framing of
310   educational internet safety initiatives is how to talk about and approach particular risks with students. Each
311   risk online should be approached and talked uniquely and originally, avoiding "one-size-fit-all" narratives
312   that assume that all risks and the nature of such risks are the same. For example, the above findings show that
313   the risks associated with certain online messaging apps are not all-encompassing, but rather narrowly
314   associated with the inappropriate use of photos. Furthermore, chatting with strangers on social media has an
315   array of potential associated risks, whereas the single risk significantly associated with chatting with strangers
316   while gaming is strangers wanting to meet in-person. Understanding this connection and the association of
317   certain platforms, actions, and risks can help inform education programming to give a more direct educational
318   lesson to students to help with improving their self-regulated behaviors, necessary skills, and understanding
319   of the potential risks around them. For educational organizations and schools, these findings should be taken
320   into consideration when building lessons and their contents; learning examples and worksheets can better
321   reflect certain risks and their association with certain behaviors. Additionally, these should inform the way
322   targeted programming for those at-risk should be handled. This data shows that girls are disproportionately
323   exposed to risky or uncomfortable situations, therefore calling on these prevention programs to address these
324   concerns with catered content. Simple changes to existing curricula might include the changing of the subject
325   in particular examples to reflect a more accurate depiction of the research above, or more large-scale changes
326   might be to offer further educational programming for those populations that are shown to be at higher risk
327   of certain exposures.

328   With the continued increase of youth internet use, the need for reflective, impactful internet safety
329   education is becoming more crucial. Schools are the main means by which a society can address digital safety

330 and citizenship issues for all children within a structured learning environment. To achieve optimal success
331 in this area, it requires teacher training and curriculum materials that are age and gender appropriate as well
332 as more research on the factors associated with online vulnerability.
333   In considering the findings of this study, several limitations are acknowledged. Since our data were cross-
334 sectional, we lack information on the temporal ordering of our variables, and we cannot assume causality. As
335 a non-random convenience sample, it is important to acknowledge that these results are not necessarily
336 generalizable outside of the sample, and that selection bias may have also influenced our reported results.
337 Furthermore, as some of the exposures and outcomes could be perceived as negative, there is potential for
338 social desirability bias to influence students' responses.
339   Its limitations notwithstanding, our results provide critical information to practitioners involved in the
340 development of educational initiatives by building a profile of potentially risky online behaviors, as well as
341 allow them to tailor their initiatives to meet the needs of more vulnerable populations. The results presented
342 in this manuscript also provide a platform for future longitudinal research to further investigate the magnitude
343 of risk associated with these behaviors. We note that the measurement of online risk has largely escaped
344 empirical investigation – to our knowledge, there is no extant screening tool for exposure to online risk that
345 has had its psychometric properties evaluated – future research should aim to address this gap. Our results
346 should be understood under the growing understanding that exposure to unsafe material online should be
347 understood as a matter of if, not when, and as such, we recommend that future research should investigate
348 how individuals engage with risky online content and the psychological processes which may influence
349 exposure.

355 **Author Contributions**: E.S. Developed the study design, drafted the introduction, and provided feedback on the
356 discussion and conclusions of the manuscript. N.H. and M.S. Performed the statistical analyses and drafted the methods,
357 results, and tables. N.S. and T.C. drafted the discussion and conclusions. All authors have read and agreed to the published
358 version of the manuscript.

367 **References**

368

369 1. George, M. J.; Jensen, M. R.; Russell, M. A.; Gassman-Pines, A.; Copeland, W. E.; Hoyle, R. H.; Odgers, C. L.,
370   Young Adolescents' Digital Technology Use, Perceived Impairments, and Well-Being in a Representative
371   Sample. *J Pediatr* **2020,** *219*, 180-187.
372 2. UNICEF *Guidelines for Industry on Child Online Protection*; UNICEF - Office of Research: 2015.
373 3. Gámez-Guadix, M.; Borrajo, E.; Almendros, C., Risky online behaviors among adolescents: Longitudinal
374   relations among problematic Internet use, cyberbullying perpetration, and meeting strangers online. *J Behav*
375   *Addict* **2016,** *5* (1), 100-107.
376 4. UNICEF *Growing up in a connected world*; UNICEF Office of Research: 2019.
377 5. Green, N., Outwardly Mobile: Young People and Mobile Technologies: The Social Context of Personal
378   Communication Technology. 2003; pp 201-217.
379 6. Ling, R., *The Mobile Connection: The Cell Phone's Impact on Society*. Morgan Kaufmann Publishers Inc.: 2004.

7. Geser, H., Pre-teen cell phone adoption: Consequences for later patterns of phone usage and involvement. In *Sociology of the Mobile Phone,* Sociology in Switzerland Online Publications Zurich: 2006.

8. UNICEF *The state of the world's children 2017. Children in a digital world.*; United Nations Children's Fund: 2017.

9. Schilder, J. D.; Brusselaers, M. B.; Bogaerts, S., The Effectiveness of an Intervention to Promote Awareness and Reduce Online Risk Behavior in Early Adolescence. *J Youth Adolesc* **2016,** *45* (2), 286-300.

10. Nesi, J.; Prinstein, M. J., Using Social Media for Social Comparison and Feedback-Seeking: Gender and Popularity Moderate Associations with Depressive Symptoms. *J Abnorm Child Psychol* **2015,** *43* (8), 1427-1438.

11. Barry, C. T.; Sidoti, C. L.; Briggs, S. M.; Reiter, S. R.; Lindsey, R. A., Adolescent social media use and mental health from adolescent and parent perspectives. *Journal of Adolescence* **2017,** *61*, 1-11.

12. Casey, B. J., Beyond simple models of self-control to circuit-based accounts of adolescent behavior. *Annu Rev Psychol* **2015,** *66*, 295-319.

13. Shulman, E. P.; Smith, A. R.; Silva, K.; Icenogle, G.; Duell, N.; Chein, J.; Steinberg, L., The dual systems model: Review, reappraisal, and reaffirmation. *Dev Cogn Neurosci* **2016,** *17*, 103-17.

14. Seabrook, E. M.; Kern, M. L.; Rickard, N. S., Social Networking Sites, Depression, and Anxiety: A Systematic Review. *JMIR Ment Health* **2016,** *3* (4), e50.

15. Spies Shapiro, L. A.; Margolin, G., Growing up wired: social networking sites and adolescent psychosocial development. *Clin Child Fam Psychol Rev* **2014,** *17* (1), 1-18.

16. Subrahmanyam, K.; Greenfield, P., Online communication and adolescent relationships. *Future Child* **2008,** *18* (1), 119-46.

17. Underwood, M. K.; Ehrenreich, S. E., The power and the pain of adolescents' digital communication: Cyber victimization and the perils of lurking. *Am Psychol* **2017,** *72* (2), 144-158.

18. Valkenburg, P. M.; Peter, J., Online communication among adolescents: an integrated model of its attraction, opportunities, and risks. *J Adolesc Health* **2011,** *48* (2), 121-7.

19. Best, P.; Manktelow, R.; Taylor, B., Online communication, social media and adolescent wellbeing: A systematic narrative review. *Children and Youth Services Review* **2014,** *41*, 27-36.

20. Sasson, H.; Mesch, G., Gender Differences in the Factors Explaining Risky Behavior Online. *J Youth Adolesc* **2016,** *45* (5), 973-85.

21. Vannucci, A.; Simpson, E. G.; Gagnon, S.; Ohannessian, C. M., Social media use and risky behaviors in adolescents: A meta-analysis. *Journal of Adolescence* **2020,** *79*, 258-274.

22. Pacheco, E.; Melhuish, N., New Zealand Children's Experiences of Online Risks and Their Perceptions of Harm. Evidence From Ngā Taiohi Matihiko O Aotearoa – New Zealand Kids Online. *SSRN Electronic Journal* **2020**.

23. Ofcom *Children and parents: media use and attitudes report 2017*; Ofcom - Research and Data Department: 2017.

24. World Medical Association Declaration of Helsinki: ethical principles for medical research involving human subjects. *Jama* **2013,** *310* (20), 2191-4.

25. Hosmer Jr, D. W.; Lemeshow, S.; Sturdivant, R. X., *Applied logistic regression*. John Wiley & Sons: 2013; Vol. 398.

26. Pregibon, D., Goodness of Link Tests for Generalized Linear Models. *Journal of the Royal Statistical Society. Series C (Applied Statistics)* **1980,** *29* (1), 15-14.

27. Suler, J., The Online Disinhibition Effect. *CyberPsychology & Behavior* **2004,** *7* (3), 321-326.

28. Palmieri, M.; Shortland, N.; McGarry, P., Personality and online deviance: The role of reinforcement sensitivity theory in cybercrime. *Computers in Human Behavior* **2021,** *120*, 106745.

29. Harriman, N.; Shortland, N.; Su, M.; Cote, T.; Testa, M. A.; Savoia, E., Youth Exposure to Hate in the Online Space: An Exploratory Analysis. *International Journal of Environmental Research and Public Health* **2020,** *17* (22), 8531.

30. Youngs, D., How does crime pay? The differentiation of criminal specialisms by fundamental incentive. *Journal of Investigative Psychology and Offender Profiling* **2006,** *3* (1), 1-19.

31. Bandura, A., *Social foundations of thought and action: A social cognitive theory*. Prentice-Hall, Inc: Englewood Cliffs, NJ, US, 1986; p xiii, 617-xiii, 617.

32. Schoon, I., Risk and Resilience. Adaptations in Changing Times. *Risk and Resilience: Adaptations in Changing Times* **2006**, 1-222.

Table 1 – Students' Social media use

Chat with strangers on social media	N (%)	Connections know in person	N (%)
I don't have a social media account	57 (8%)	I don't have a social media account	82 (11%)
Never	222 (30%)	I am not sure how many	130 (18%)
I did it a few times	215 (29%)	Some of them	149 (20%)
Sometimes	154 (21%)	Most of them	272 (37%)
Often	85 (12%)	All of them	100 (14%)
Share personal information on social media	**N (%)**	**Friends chat with strangers on social media**	**N (%)**
I don't have a social media account	76 (10%)	I don't know	195 (27%)
Never	433 (59%)	Never	84 (11%)
Sometimes	191 (26%)	Sometimes	319 (44%)
Often	33 (5%)	Often	135 (18%)
Talked to about online safety	**N (%)**	**Use private messaging more than texting**	**N (%)**
Never	35 (5%)	No	581 (79%)
Sometimes	280 (38%)	Yes	78 (11%)
Often	418 (57%)	Not sure	74 (10%)
Trusted adult	**N (%)**	**Chat with strangers while gaming**	**N (%)**
No	79 (11%)	I do not play video games	97 (13%)
Yes	507 (69%)	Never	260 (36%)
Not sure	121 (16%)	Sometimes	215 (29%)
It depends on the situation	26 (4%)	Often	161 (22%)
Presence of parental controls on computer or other devices	**N (%)**	**Play video games**	**N (%)**
No	378 (52%)	Never	163 (22%)
Yes	236 (32%)	Often	289 (40%)
Not sure	119 (16%)	Sometimes	281 (38%)
Use Kik	**N (%)**	**Number of social media platforms used daily**	**N (%)**
No	696 (95%)	0	112 (15%)
Yes	37 (5%)	1	87 (12%)
Use Telegram	**N (%)**	2	165 (23%)
No	724 (99%)	3	188 (26%)
Yes	9 (1%)	4	111 (15%)
Use WhatsApp	**N (%)**	5	52 (7%)
No	666 (91%)	6 or more	18 (2%)
Yes	67 (9%)		

Table 2 – Unsafe Online Situations

Situation	N (%)
Bullying or harassment by friends or acquaintances	214 (29%)
Getting involved in unwanted conversations in a chat room, social networking site or on email	187 (26%)
Coming across sexual images or content	224 (31%)
Coming across images of violence	236 (32%)
Someone trying to sell me drugs or alcohol	101 (14%)
Someone using my photos in an inappropriate way	42 (6%)
A stranger trying to meet with me	125 (17%)
Coming across hate groups trying to convince me of their views	102 (14%)
Number of Situations Encountered	**N (%)**
0	291 (40%)
1	147 (20%)
2	87 (12%)
3	86 (12%)
4	45 (6%)
5	34 (5%)
6	17 (2%)
7	8 (1%)
8	18 (2%)
Mean (SD)	1.7 (2)
Median (Range)	1 (0-8)

Table 3 – Multiple Logistic Regression Models

Covariate	OS 1 Bullying	OS 2 Unwanted conversations	OS 3 Sexual images	OS 4 Violent images	OS 5 Drugs & alcohol	OS 6 Photos used inappropriately	OS 7 Stranger	OS 8 Hate Groups	At least one unsafe event
Continuous and Dichotomous Predictors – OR (95% CI)									
Age (15 and older vs. 14)	1.1 (0.8-1.6)	1.2 (0.8-1.8)	1.2 (0.8-1.7)	1.7 (1.2-2.4)**	1.4 (0.8-2.2)	0.8 (0.4-1.6)	1.5 (0.9-2.2)	1.2 (0.8-2)	1.3 (0.9-1.8)
Gender (Female vs. Male)	1.7 (1.2-2.4)**	2 (1.4-2.9)**	2 (1.4-2.9)**	1.9 (1.3-2.7)**	1.4 (0.9-2.3)	1.8 (0.9-3.9)	1.6 (1-2.7)	1.6 (1-2.5)	2 (1.4-2.7)**
Race (White vs. Non-White)	1.3 (0.9-1.9)	1.1 (0.8-1.7)	1.5 (1-2.1)*	1.1 (0.8-1.6)	1.2 (0.7-1.9)	0.9 (0.4-1.9)	0.7 (0.5-1.1)	0.8 (0.5-1.3)	1.2 (0.9-1.7)
Grades (B and lower vs. B+ and higher)	-	-	-	1.6 (1.1-2.3)*	-	-	-	-	-
Kik Use (Yes vs. No)	1.8 (0.9-3.9)		1.9 (0.9-4)		1.2 (0.5-3)	3.1 (1.1-9.2)*	1.6 (0.7-3.8)	2.1 (0.9-4.9)	1.4 (0.6-3.2)
WhatsApp Use (Yes vs. No)	-	-	-	-	-	-	1.2 (0.6-2.3)	-	1.2 (0.6-2.2)
Telegram Use (Yes vs. No)	-	-	-	-	-	-	-	-	-
Chatting with strangers on social media (Any Frequency of Communication vs. Never/ I don't have a social media account)	1.9 (1.3-2.8)**	1.8 (1.2-2.9)*	1.7 (1.1-2.5)*	1.7 (1.2-2.6)*	1.7 (1-3.1)	4.5 (1.5-13.3)*	2.1 (1.2-3.4)*	1.8 (1-3.2)*	2.3 (1.6-3.4)**
Sharing Personal Information on Social Media (Any Frequency of Sharing vs. Never/ I don't have a social media account)	-	1.2 (0.8-1.8)	-	-	-	1.7 (0.8-3.5)	1.3 (0.8-2)	-	1.3 (0.9-1.9)
Playing video games (Any Frequency of Use vs. Never)	-	-	-	-	-	-	0.9 (0.5-1.6)	-	-
Chatting with strangers while playing video games (Any Frequency of Communication vs. Never/ I don't play video games)	-	-	-	-	-	-	0.5 (0.3-0.8)*	-	-
Having someone talk to you about online safety (continuous)	-	-	-	1.2 (0.9-1.6)	-	-	-	-	1.2 (0.9-1.6)
Social Media Use (continuous)	1.14 (1.01-1.28)*	1.07 (0.93-1.24)	1.11 (0.97-1.27)	1.24 (1.08-1.42)**	1.13 (0.95-1.35)	1.11 (0.86-1.43)	1.01 (0.87-1.18)	1.02 (0.86-1.22)	1.16 (1.01-1.32)*
Categorical Predictors with 3+ Categories - Chi-Squared (df); *p*-value									
Use Private Messaging Apps More than Texting (categorical)	-	-	-	-	-	-	-	-	-
Amount of followers known in person (categorical)	-	3.52 (4); *p*=0.474	3.14 (4); *p*=0.534	7.24 (4); *p*=0.124	5.57 (4); *p*=0.216	-	-	5.28 (4); *p*=0.26	1.08 (4); *p*=0.897
Frequency that friends chat with strangers on social media (categorical)	-	7.5 (3); *p*=0.056	10.43); *p*=0.015	9.45 (3); *p*=0.024	6.75 (3); *p*=0.077	-	6.83 (3); *p*=0.078	10.7 (3); *p*=0.014	11.8 (3); *p*=0.008
Having a trusted adult to ask for help with unsafe online situations (categorical)	6.78 (3); *p*=0.079	3.91 (3); *p*=0.271	-	-	-	4.35(3); *p*=0.226	-	-	-

Presence of parental controls on your computer (categorical)	-	-	-	-	-	-	-	-	-
Frequency that friends chat with strangers on social media	Pairwise comparison of marginal linear predictors (if applicable) – OR (95% CI)							
"Never" vs "I don't know"	-	-	1.8 (0.9-3.4)	1.1 (0.6-2.2)	-	-	-	1.1 (0.4-2.9)	0.9 (0.5-1.6)
"Sometimes" vs "I don't know"	-	-	2.1 (1.3-3.3)**	1.5 (1-2.3)	-	-	-	1.2 (0.6-2.2)	1.8 (1.2-2.8)**
"Often" vs "I don't know"	-	-	1.5 (0.9-2.7)	2.2 (1.3-3.7)**	-	-	-	2.6 (1.3-5)*	1.6 (1-2.7)
"Sometimes" vs "Never"	-	-	1.2 (0.6-2.2)	1.3 (0.7-2.5)	-	-	-	1.1 (0.4-2.9)	2 (1.1-3.5)*
"Often" vs "Never"	-	-	0.9 (0.4-1.8)	2 (1-4)	-	-	-	2.4 (0.9-6.6)	1.7 (0.9-3.3)
"Often" vs "Sometimes"	-	-	0.7 (0.5-1.2)	1.5 (1-2.3)	-	-	-	2.2 (1.3-3.7)*	0.9 (0.5-1.4)

* *p* < 0.05

** *p* < 0.01

OS 1 – Bullying or harassment by friends or acquaintances

OS 2 – Getting involved in unwanted conversations in a chat room, social networking site or on email

OS 3 – Coming across sexual images or content

OS 4 – Coming across images of violence

OS 5 – Someone trying to sell me drugs or alcohol

OS 6 – Someone using my photos in an inappropriate way

OS 7 – A stranger trying to meet with me

OS 8 – Coming across hate groups trying to convince me of their views

Table 4 – Negative Binomial Model for Counts of Types of Unsafe Online Situations

Covariate	IRR (95% CI)
Age (15 and older vs. 14)	1.1 (1.0-1.4)
Gender (Female vs. Male)	1.5 (1.3-1.9)**
Race (White vs. Non-White)	1.1 (0.9-1.3)
Kik Use (Yes vs. No)	1.2 (0.8-1.8)
Chatting with strangers on social media (Any Frequency of Communication vs. Never/ I don't have a social media account)	1.6 (1.3-2.0)**
Social Media Use (continuous)	1.08 (1.01-1.15)*
	Chi-Squared (df); *p*-value
Frequency that friends chat with strangers on social media (categorical)	14.6 (3) *p*=0.002
Frequency that friends chat with strangers on social media	**Pairwise comparison of marginal linear predictors (if applicable) – IRR (95% CI)**
"Never" vs "I don't know"	1.3 (0.9-1.8)
"Sometimes" vs "I don't know"	1.5 (1.2-1.9)**
"Often" vs "I don't know"	1.6 (1.2-2.1)**
"Sometimes" vs "Never"	1.2 (0.8-1.6)
"Often" vs "Never"	1.2 (0.9-1.8)
"Often" vs "Sometimes"	1.1 (0.8-1.3)

* *p*<0.05

** *p*<0.01