



ELSEVIER

Contents lists available at ScienceDirect

Computers &amp; Education

journal homepage: [www.elsevier.com/locate/compedu](http://www.elsevier.com/locate/compedu)

# Identity profile of young people experiencing a sense of risk on the internet: A data mining application of decision tree with CHAID algorithm

José Manuel Muñoz-Rodríguez<sup>a</sup>, Carmen Patino Alonso<sup>b</sup>, Teresa Pessoa<sup>c</sup>,  
Judith Martín-Lucas<sup>a,\*</sup>

<sup>a</sup> Faculty of Education, University of Salamanca. Department of Theory and History of Education. Paseo de Canalejas, 169, 37008. Salamanca, Spain

<sup>b</sup> Statistics Department. University of Salamanca. Calle Alfonso X el Sabio, s/n, 37007. Salamanca, Spain

<sup>c</sup> Faculty of Psychology and Education Sciences, University of Coimbra, Rua do Colégio Novo, 1 3000-115, Coimbra, Portugal

## ARTICLE INFO

### Keywords:

Risk  
Perception  
Internet  
Identity  
Adolescents  
CHAID

## ABSTRACT

Social media, interactive games, and digital platforms are just some of the ways in which the internet has become part of our lives, and more specifically so among young people. Among other things, this technology is the reason young people are spending increasingly more time online than offline, prompting numerous risks that they are sometimes unaware of. The following quantitative research is based on this premise, and it pursues two objectives, namely, to discover whether or not young people experience a sense of risk linked to their use of the internet and identify the kind of young people that say they do or do not perceive any such risk. A total of 1991 young people answered the *Technology Use Studies* questionnaire. The classification tree technique based on the Chi-square Automatic Interaction Detector (CHAID) was applied to identify the predictive variables associated with the sense of risk in young people. Two highly significant results have been forthcoming: firstly, we present the factors with the biggest impact on a young person's internet-related sense of risk, involving isolation when they cannot go online, followed by variables such as gender, the use of TikTok, and videogames, the consumption of pornography, and distraction from study time. Secondly, our results indicate that the young people that feel isolated when they cannot go online are precisely the ones that do not experience a sense of online risk. It is therefore of vital importance for education to generate and provide young people with strategies that allow them to establish a healthy and responsible relationship with technologies such as the internet.

## 1. Introduction

Today, the different environments provided by digital platforms, social media, and interactive games, among many others, mean that the screens on smartphone, tablets, and computers, for example, have become the daily setting for young people's lives, where they develop their identities and learn. No one now denies that the virtual environment has become part of our lives, and more

\* Corresponding author.

E-mail addresses: [pepema@usal.es](mailto:pepema@usal.es) (J.M. Muñoz-Rodríguez), [carpatino@usal.es](mailto:carpatino@usal.es) (C. Patino Alonso), [tpessoa@fpce.uc.pt](mailto:tpessoa@fpce.uc.pt) (T. Pessoa), [judithmartin@usal.es](mailto:judithmartin@usal.es) (J. Martín-Lucas).

<https://doi.org/10.1016/j.compedu.2023.104743>

Received 14 November 2021; Received in revised form 11 January 2023; Accepted 18 January 2023

Available online 28 January 2023

0360-1315/© 2023 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

specifically of young people's lives (Spain's National Office of Statistics - INE, 2020; Telefonica Foundation, 2020). The growth of the internet has brought about a new form of social organisation that has introduced new ways of socialising and communicating at both individual and social level. Young people's traditional areas of shared existence are receding, and a new scenario for relationships is appearing that sometimes takes over and even replaces other traditional settings for interaction. Digital technology has prompted not just the migration from physical to virtual space, but also the creation of new scenarios in which there is life, as settings for social relationship and actions (Pangrazio, 2019; Sánchez-Rojo, García del Dujo, Muñoz-Rodríguez, & Dacosta, 2022). They cater for different kinds of self-expression, of manifesting the "self" that we see in ourselves – the one we want other people to see or the one we think we can see. In other words, the internet and social media amplify social relationships and the options for building an online identity. This even favours the construction of an online intelligence by introducing new mechanisms of social recognition that are not always positive (Malone & Bernstein, 2015). Young people's identity practices have become technologically mediated actions; digital times and spaces play a significant part in young people's social lives (Astleitner, Bains, & Hörmann, 2023). This is true to such an extent that the current literature stresses that the traditional boundary between the online and offline worlds is becoming blurred (Resnik & Bellmore, 2019), fading, and this is now more accurately described as *onlife* (Floridi, 2015).

Within this scenario, there are numerous social players and educational institutions that have expressed their concern (Gür & Kiliç Türel, 2022; Hodkinson, 2017; Torrijos Fincias, Serrate González, Martín Lucas, & Muñoz Rodríguez, 2021; Yau & Reich, 2019), as adolescence is a critical time in the formation of an identity (Kim, 2021). Families, educators, and teachers, those that oversee regulations and legislation in schools and/or local authorities, do not know how to respond to the numerous situations that are constantly appearing. They are calling both for diagnoses of the situation and for the adoption of measures, strategies, and even rules for tackling an issue that, much more often than we would like, is considered an actual problematic. Problems, in short, that impact upon human lives, not on abstract entities, but instead on young people and groups with names and faces, whose mechanisms for shaping their identities are adopting untried and untested approaches, which often involve risks and drawbacks (Calvete, Fernández-González, González-Cabrera, Machimbarrena, & Orue, 2020; Soler, López-Sánchez, & Lacave, 2018; Plaza de la Hoz, 2021).

Prior studies have already sounded the alarm (Valdemoros, Sanz, & Ponce de León, 2017; Vannucci, Simpson, Gagnon, & Ohanessian, 2020): young people are turning their backs on their family haven at an earlier age and in a more drastic manner; the greater their digital consumption, the weaker their emotional ties with their family environments; the longer they spend online, the less they think critically and self-reflect upon their social behaviour, and the greater their obsession with self-expression linked purely and simply to pleasure and entertainment, heightening their individualism and leading to a series of risks of which they are often not even aware (Mace, 2020).

Most of the studies thus far have focused on identifying and quantifying the risks associated with internet use, screens, and social media (e.g., Murciano Hueso, Gutiérrez Pérez, Martín Lucas, & Huete, 2022; Benchea, 2021; Marzilli, Cerniglia, Ballarotto, & Cimino, 2020; Yildiz Durak & Saritepeci, 2019; Zhou & Zhang, 2019) or found that young people perceive a certain feeling of risk associated with their use of the internet (Gaspard, 2020). Nevertheless, there are no studies that seek to identify the profile of young people that perceive the risk that the use and consumption of this type of technology entails. The originality of our contribution therefore involves identifying the profile of those young people that do not perceive the risks inherent to the use of technologies such as the internet.

Faced with this phenomenon, we need studies that use young people's own words to diagnose the sense of risk they actually perceive and the extent and nature of those risks in order to go beyond the pedagogy of limits and incorporate constructive values. This means there is a need to identify the profile of those young people using the internet that do not perceive its risks, as it is vitally important to design educational measures to suit this type of profile. Such is the premise that informs this research, whose purpose is to understand young people's sense of risk and discover whether there is a profile that enables us to identify those that are oblivious to online risks in order to obtain a channel for proposing possible ways of addressing them through education. The study that now follows seeks to answer the following two questions: do young people experience a sense of risk? If so, what are the common features shared by those young people that do and do not perceive online risks? By first answering these questions, we can pinpoint the sense of risk associated with the use of the internet that might be affecting their identity building, and secondly, we may lay the foundations for rolling out quality education strategies that enable us to anticipate and act upon the profile of young people that may be more vulnerable to the risks involved in internet use.

Before continuing, it should be clarified that we understand a sense of risk to mean the risk that young people experience when they go online. In addition, when we speak of internet use, we do so in general terms, referring to the use of this technology in all its manifestations, including social media, interactive games, and digital platforms, for example.

## 2. Theoretical underpinnings and literature review

This research's theoretical underpinnings are described from three angles, which in turn conflate in a pedagogical model that explains their meaning. On the one hand, we base ourselves on the theory that the use of the internet maintains two basic dimensions that come together to inform human identity: space and time (Mace, 2020; Pangrazio, 2019). Without that spatiotemporal dimension of human activity on the internet, we could not speak of education, of human development. It is not an aspatial place nor an atemporal one, and much less so a mere tool, but instead the time and place in which young people occupy and act, relate to each other, move, and communicate; all of which are primary development processes (García del Dujo, Vlieghe, Muñoz-Rodríguez, & Martín-Lucas, 2021). Studying the possible sense of risk in these identity-building processes enables us to confirm this theory and, in turn, extend its rationale whereby we can understand technology as culture, inasmuch as it creates new ways of thinking, acting, and being. These are alternative ways of acting and being in the world, impacting upon the reontologisation of human beings, and in this case, of youth (Vlieghe & Zamojski, 2019).

On the other hand, these identity rebuilding processes are faced considering the virtual environment to be an everyday setting for coexisting, acting, and relating (Gui, Gerosa, Argentin, & y Losi, 2023). Until only recently, studies have addressed the contrast between online and offline lives, in the digital and real worlds, respectively (Memon, Sharma, Mohite, & Jain, 2018). We base this research on the theory defined by Floridi (2015), amongst others, which in terms of young people refers to online experience, living, and identity. A large part of a young person's life takes place online (Al-Furaih & Al-Awidi, 2021; Muñoz-Rodríguez, Torrijos, Serrate, & Murciano, 2020; Floridi, 2015). The data we present corroborate this notion that transports us to an ephemeral, unstable world, at risk of building an identity on quicksand due to the lack of a sense of risk that the use of technologies such as the internet entails (Gaspard, 2020).

Finally, and as already noted, we base ourselves on a young person's status as a non-digital native or immigrant. Ever since Prensky (2001) coined the phrase *digital native*, the bulk of the research has been using this concept to explain the social and educational changes brought about by technology. In our case, however, we formulate the hypotheses that young people are not digital natives, based on their lack of critical thinking and feeling when acting and being online (Muñoz-Rodríguez, Dacosta, & Martín-Lucas, 2021). This study's findings specifically confirm this hypothesis.

These three postulates in conjunction inform the need to base the studies on education and technology through parameters close to post-critical pedagogy. Over and above criticising young people for their common uses and misuses of technological environments, we should reappraise education in terms of positive and edifying pedagogies. "... maybe the time has come to question the assumption that these characteristics are always inherently negative" (Hodgson, Vlieghe, & Zamojski, 2020, p. 8).

Furthermore, we consider that all educational processes involve identity building, regardless of a person's age. In other words, education means forming individuals, which in turn means building both personal and collective identities. An identity-building process involving young people is the mainstay of all such processes, it is education itself; and young people constitute the group most affected by technological mediation and the sense of risk in this selfsame mediation. The reason for this is that the computer, the screen, is where this identity is built, where a young person's education is informed. The aim is to address a young person's identity in interaction with the computer and the possible risks they perceive from an essentially dynamic understanding of identity (Bernal, 2021), with a view to better and greater self-determination in front of the screen, the development of a more constructive identity; in other words, a more enlightening education (Thoilliez, 2020).

In line with the above, this research provides new insights into the concept of converting our pedagogical action into a suspicion and a need to uncover the truth, and not criticise it beforehand.

As we have already noted, technology and virtual space have ushered in new configurations, not only in spatiotemporal terms, but also in the field of human communication and social media. They have prompted a huge development in the world of leisure and informal socialising, incorporating numerous benefits within the educational, professional and social ambits (Heidari et al., 2020). Further still, these technologies have permitted social relations and the daily work and activities of the vast majority of the population during the lockdown imposed by Covid-19. Nevertheless, it cannot be ignored that these technologies reshape us as users, affecting our ways of being, thinking and living (García del Dujo, Vlieghe, Muñoz-Rodríguez, & Martín-Lucas, 2021; Sánchez-Rojo & Martín-Lucas, 2021). While technological progress has provided major gains in accessing, managing and sharing information, among other aspects, it has, by contrast, posed a series of hazards for those people that fail to abide by certain clearly defined rules. The risks involved mainly affect young people (Yildiz Durak & Saritepeci, 2019), as they precisely constitute the most vulnerable social group because they are immersed in a period of psychosocial change during a time when the greatest use is made of this technology. In Spain's case, the use of the internet, social media, and videogames is a common practice for 99.9% of young men and 99.6% of young women aged between 16 and 24 (INE, 2020). This same report reveals that over 90% of adolescents aged 14 and 15 have their own mobile phone. In short, young people can go online whenever they like and wherever they are, spending their time on a raft of different activities, with the main ones being the chance for interaction with their peer group and self-expression through social media, carrying out academic or educational tasks, or simply as a way of occupying their free time and combatting boredom (Muñoz-Rodríguez, Torrijos, Serrate, & Murciano, 2020; Murciano Hueso, Gutiérrez Pérez, Martín Lucas, & Huete, 2022; Brough, Literat, & Ikin, 2020; Liu, Wright, & Hu, 2018).

Over the past years, we have relied on young people's assumed digital nativity, as identified by Prensky (2001), to explain that young people know how to use and manage this technology. The reality, however, appears to be quite different; young people do not have the ability to adopt a critical and responsible approach to the use of technological devices (Escofet, López, & Álvarez, 2014; Warden, Yi-Shun, Stanworth, & Chen, 2020). There is even a systematic study on this digital nativity that concludes that young people do not actually have this competence, mainly because they lack a critical sense (Kirschner & De Bruyckere, 2017). As we were questioning this digital nativity, stress was placed on highlighting the pedagogy of limits, associated with restricting the time spent online and the sites accessed, being equally aware that this was insufficient. The risks also lay – and indeed still lie – in the little or large amount of time a young person is allowed to spend online, as well as in those sites that do not initially appear to pose any kind of risk, until they do. Nevertheless, there are few studies on whether young people experience a sense of risk associated with their use of the internet and the variables that might affect this perception. For example, a study conducted by Gaspard (2020) has revealed that 39.5% of boys and 45.5% girls consider that internet use involves more concerns than opportunities. Most of the risks often have negative repercussions that on a socioemotional level appear in young people as a result of the misuse of social media and online games (Kim, 2021; Zhou & Zhang, 2019). These addictions are sometimes the cause, and at others the consequence, of a feeling of social isolation (Benchea, 2021; Chandra, Mondal, Ansary, & Saha, 2018; Firat, Karakaya, & Hergul, 2011) and end up prompting problems of stress and anxiety when the young person is not online (Marzilli et al., 2020). On average, young people spend 3 h a day on social media (Savoia, Harriman, Su, Cote, & Shortland, 2021), which sometimes becomes a source of distraction for their studies, and even reduces the time they have to enjoy quality relations with their peer group (Plaza de la Hoz, 2021; Hu & Yu, 2021), and may also end up

interfering with their family relationships (Marzilli et al., 2020). What's more, there are studies that link videogames to the consumption of pornography and the compulsive use of online communities (Escario & Wilkinson, 2020), often leading to internet addiction (Casaló & Escario, 2019). The concept of *alone together* coined by Turkle (2011) now acquires greater meaning than ever, as more and more young people are learning to be alone without actually being so. It is no surprise that some of these risks and issues arising from internet use already have a name of their own, such as *phubbing*, referring to those that pay more attention to their technological devices than to the people around them (Álvarez & Moral, 2020), and *FOMO (Fear of Missing Out)*, which refers to the fear of not being online and, therefore, unable to connect to social media (Al-Furaih & Al-Awidi, 2021; Astleitner et al., 2023). In sum, these risks trigger a sense of isolation when an individual is not connected to the internet and their social media.

Unfortunately, most of these risks do not come alone, but instead appear in clusters (Calvete et al., 2020). Furthermore, there are a number of studies that find gender differences regarding exposure to risk on the internet and the way it is perceived, with young women being more likely than their male counterparts to be victims of cyberbullying (Byrne, 2021; Evangelio, Rodríguez-González, Fernández-Río, & González-Villora, 2022). In turn, young men record inappropriate behaviour associated with the consumption of pornography (Chen, Chang, Wang, Wang, & Wei, 2021) or excessive use of videogames (Gómez Galán, Lázaro-Pérez, & Martínez-López, 2021; Masanet, Pires, & Gómez-Puertas, 2021).

Among these risks, the one that has thus far been most widely studied and has had the greatest impact is cyberbullying (Çevik, ATA, & Çevik, 2021) associated with emotional and psychological problems of socialisation, and which is increasingly recording ever more dramatic cases (Aboujaoude, Savage, Starcevic, & Salame, 2015; Evangelio et al., 2022; Garaigordobil & Larrain, 2020; Li, Wu, & Hesketh, 2023; Núñez, Álvarez-García, & Pérez-Fuentes, 2021; Selwyn & Aagaard, 2020)

The following study seeks to go beyond the identification and quantification of the risks associated with internet use, its causes and consequences, to focus on exploring the profile of young people with an internet-related sense of risk. Our main interest, and the reason for the originality of this research, is to verify whether those young people that do not experience an internet-related sense of risk share any common characteristics, and therefore provide the basis for an educational programme.

### 3. Method

A quantitative study was conducted within the framework of a non-experimental ex-post facto design because the variables were not intentionally adjusted. In line with prior studies conducted by our research group (Muñoz-Rodríguez, Torrijos, Serrate, & Murciano, 2020), the aim is to further develop an analysis of young people's identity profile, on this occasion considering the variables associated with the sense of risk this cohort perceives online.

#### 3.1. Participants

Sample size was calculated based on the universe approach, as it should be remembered that the population under study is almost entirely at school, whereby any representative sample should gather data mainly on the population within a school context. According to official data for the 2018–2019 school year provided by Spain's Ministry of Education, Culture and Sport, the net rates of schooling for the Spanish population are as follows: almost 100% for those aged between 12 and 15; 96% for those aged 16; 89.8% for 17-year-olds, and 79.5% for 18-year-olds. The response rate of this questionnaire was 89.8%.

The sample consisted of 1991 individuals, of whom 57.2% were female and 42.8% male, aged between 12 and 18, from fifteen of Spain's autonomous communities or regions (24.4% Castilla y León, 21.2% Castilla la Mancha, 10.6% Catalonia, 9.7% Balearic Isles, 8.3% Madrid, 8.1% Aragon, 7.1% Cantabria, 5.3% Andalusia, 2% Extremadura, 2% Community of Valencia, 0.6% the Basque Country, 0.3% Galicia, 0.2% Principality of Asturias, and 0.1% La Rioja and the Canary Islands, respectively). Specifically, contact was made with 31 secondary schools nationwide segmented by the size of their location (rural or urban) and socioeconomic environment, with 31.2% of the sample living in the countryside and 68.4% in a city.

#### 3.2. Instrument

Use was made of the questionnaire *Estudios sobre el uso de tecnología* [Studies on the use of technology]. After a pilot-test, a 19-item final questionnaire was administered, obtaining a satisfactory index for internal consistency (Cronbach's Alpha = 0.713). This anonymous questionnaire is divided into the following four sections: a) sociodemographic profile (3 items), b) academic record (1 item), c) online scenarios and usages (13 items), and d) results and effects (5 items).

The data were gathered online through the tool *Google Form*, which is the method most suited to this kind of study, not only because of the social and health situation caused by Covid-19 but also considering the subject-matter and the target population (Niemi, Portney, & King, 2008).

#### 3.3. Procedure

The data were gathered between September 2020 and January 2021. The teachers and heads at secondary schools nationwide were contacted by an email containing a letter of introduction to the research, its ethical criteria (voluntary participation and data confidentiality), and the questionnaire, together with the instructions for completing it.

### 3.4. Statistical analysis

The data analysis involved the IBM SPSS Statistics software, version 26 for Windows. The category variables were stated as the distribution of frequencies.

The methodology was articulated in two stages (Fig. 1).

Stage 1: Separate logistic regressions were performed using the inverse selection of conditional variables method to identify the factors of young people’s sense of online risk. Sense of risk was taken as the dependent variable (1 = risk, 0 = no risk). The independent or predictor variables were those related to young people’s sense of isolation and unhappiness, dissatisfaction, whether or not they had been bullied or abused, and their management of online time and scenarios (Table 1).

Stage 2: The significant predictors in the previous stage were then used in the decision tree with CHAID algorithm.

Data-mining involves extracting information from a data set, helping to identify the relevant information and expediting decision-making (Kesavaraj & Sukumaran, 2013). Classification is one of data mining’s functions, and the decision tree with CHAID algorithm is the supervised classification technique most often applied. It involves classifier with a tree-shaped structures that reflects a procedure for classifying data in which the tree’s upper node is the root node (Chahal, 2013). This research has used a data mining application of decision tree with CHAID algorithm to help to identify the type of young people that affirm whether or not they perceive a sense of risk linked to their internet use.

The variables associated with a sense of online risk in the bivariate analyses were included in the multivariate mode via a Chi-square Automatic Interaction Detector (CHAID) algorithm.

The data were split into training (75%) and test (25%) sets. A tenfold cross-validation analysis was performed as an initial evaluation of the algorithm. The original study cohort is randomly grouped into ten subsets of equal size. Nine of the ten subsets in each of these steps are used to fit the model and one subset is used as a validation set for the misclassification risk. Finally, all the risks are used to calculate an overall measure of the performance of the cross-validation model. This estimate of the misclassification risk is calculated as the average of all the risks obtained in the cross-validation process. The percentage of the overall correct classification and the risk value of misclassification for the training set and cross-validation were used to evaluate the performance of the algorithm.

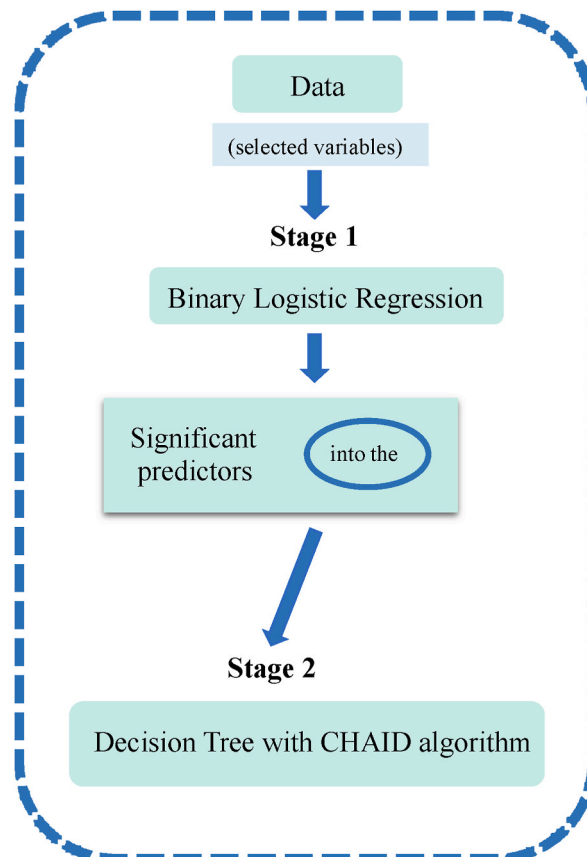


Fig. 1. Statistical analysis flow chart.

**Table 1**  
Detail of explanatory variables with the sense of risk.

| Dimension       | Description   | Coding   |
|-----------------|---|--|
| Gender          | Biological sex of participants  | 1 = Female<br>2 = Male   |
| Assessment      | Average school marks in the preceding academic year   | 1 = A<br>2 = B<br>3 = C (pass)<br>4 = D (fail)                               |
| Isolation       | I feel isolated when I'm not online   | 1 = Completely agree<br>2 = Agree<br>3 = Disagree<br>4 = Completely disagree |
| Unhappiness     | It has made me unhappy (for not doing more things offline)  | 0 = Never<br>1 = Sometimes<br>2 = Often                                      |
| Dissatisfaction | My online life has not given me much satisfaction   | 1 = Sometimes<br>2 = Often   |
| Abuse           | Someone has tried or managed to abuse me  | 0 = Never<br>1 = Sometimes<br>2 = Often                                      |
| Bullying        | I have suffered online bullying   | 1 = Sometimes<br>2 = Often   |
| Time management | I lose track of time when I'm online  |  |
| Scenarios       | YouTube<br>WhatsApp<br>Instagram<br>TikTok<br>Facebook<br>Twitter<br>Telegram<br>Snapchat<br>Music (Spotify or similar)<br>Online gaming (any)<br>Online betting (any)<br>Pornography | 0 = No use<br>1 = Limited use<br>2 = Frequent use                            |

### 3.5. Chi-square Automatic Interaction Detector (CHAID) algorithm

CHAID is an algorithm spawned from THAID, proposed by Kass (1980). It is one of the most popular supervised learning methods based on statistics for the creation of decision trees. CHAID stands for Chi-squared Automatic Interaction Detector. CHAID modelling is an exploratory data analysis method used for studying the relationships between a criterion variable (dependent) and a wide range of possible predictor variables (independent) that may interact with each other. The CHAID algorithm presents the relationships between the variable in a hierarchical manner and displays the results conveniently in the form of a classification tree that contains many uniform subsets with the splitting of multidirectional nodes informed by a set of heterogeneous data, selecting the explanatory variables that record a significant interaction with a response variable.

The process of building the decision tree with CHAID algorithm involves the steps of merging, splitting, and stopping steps. The best divisions are identified by the chi-square statistic. The CHAID algorithm therefore involves the following steps.

**Step 1.** Fusion of the categories of predictor variables: this involves merging the values of the explanatory variables that are not statistically uniform regarding the dependent variable and separating those other ones that are heterogeneous. The algorithm continues in an iterative manner until all the category pairs are treated as statistically different; in other words, if the p-value obtained exceeds a given fusion threshold, the algorithm merges certain categories without statistically significant differences. The type of predictor informs the groupings, which means that any combination is possible for a nominal variable (free predictor), while in the case of an ordinal variable (monotonous predictor), only contiguous categories may be combined, and use may also be made of a "floating predictor"; that is, one in which all the categories except one are ordinal.

**Step 2.** Best predictor: the algorithm looks for the best predictor to better differentiate the values of the dependent variable. On the resulting tree, the main independent variable appears in the first node in the classification, which displays the lowest p-value and, therefore, the one most strongly associated with the dependent variable. Each variable is assessed in terms of its association with the dependent variable according to the p-value obtained with the statistical test, whereby the algorithm selects the best predictor to form the tree's first division; that is, the explanatory variable with the closest association with the dependent variable (the one whose chi-square test provided the lowest p-value). If this value is equal to or lower than a given threshold of division established the variable is the used as a division variable for the node in question, and each one of the fused categories of the division variable defines a branch node. Once the node in question has been divided, the branch nodes are examined to see if they allow for further divisions through the application of the fusion/division process. The first two steps involve the application of the Pearson chi-square test for independence for specifying the conditions of fusion and division.

**Table 2**  
Odds ratios and 95% confidence intervals from binary logistic regression models showing the association between young people’s sense of risk and explanatory variables.

| Variable                                | OR (95% CI)         | p-value |
|---|---------------------|---------|
| <u>Gender</u>                           | 1.481 (1.239–1.239) | <0.001  |
| <u>Marks</u>                            | 0.412 (0.248–0.684) | 0.001   |
| YouTube                                 | 0.734 (0.511–1.054) | 0.094   |
| WhatsApp                                | 0.651 (0.400–1.059) | 0.084   |
| <u>Instagram</u>                        | 0.497 (0.377–0.654) | <0.001  |
| <u>TikTok</u>                           | 0.700 (0.563–0.871) | 0.001   |
| Spotify                                 | 0.460 (0.184–1.150) | 0.097   |
| <u>Twitter</u>                          | 0.717 (0.531–0.968) | 0.030   |
| Telegram                                | 0.940 (0.383–2.311) | 0.893   |
| Snapchat                                | 0.667 (0.432–1.031) | 0.068   |
| Music                                   | 0.757 (0.559–1.024) | 0.071   |
| <u>Videogames</u>                       | 0.564 (0.447–0.713) | <0.001  |
| Betting                                 | 0.731 (0.264–2.025) | 0.547   |
| <u>Pornography</u>                      | 0.456 (0.296–0.703) | <0.001  |
| Been abused                             | 1.075 (0.508–2.271) | 0.851   |
| Been bullied                            | 0.635 (0.349–1.155) | 0.137   |
| <u>Distracting from studies</u>         | 0.549 (0.422–0.714) | <0.001  |
| <u>Feeling isolated when not online</u> | 0.659 (0.537–0.810) | <0.001  |
| Feeling unhappy                         | 1.048 (0.594–1.848) | 0.872   |
| <u>Using media is a waste of time</u>   | 1.389 (1.029–1.875) | 0.032   |

**Step 3.** The first and second steps are repeated until one of the following rules for ending the algorithm has been fulfilled, and which need to be established before conducting the analysis: there is no other predictor for which a p-value is lower than a given level of significance, the highest number of levels on the classification tree has been reached, and the lowest number of observations cannot be reached in any further node (Alkhasawneh, Kalthum Ngah, Tien Tay, Ashidi Mat Isa, & Subhi Al - Batah, 2014).

The CHAID algorithm is illustrated in the following diagram Fig. 2.

CHAID algorithm is merely one of many techniques used in decision trees; its output is highly visual and easy to interpret, and no type of distribution of independent variables is assumed a priori because it relies on the use of the Chi-square statistic. One of the disadvantages of this method is that it requires large samples.

Compared to other algorithms in data mining, CHAID in decision trees follows a non-parametric approach, which means it does not require distributive assumptions such as normality and linearity. CHAID algorithm does not require the prune back operation used in, for example, the CART algorithm developed by Breiman, Friedman, Olshen, and Stone (1984). However, CHAID cannot process continuous data, so they need to be converted into categorical variables. The QUEST algorithm (Loh & Shih, 1997) can only process binary data and it involves the assumption that the target variable is a continuous one, whereby it could not be used in this study.

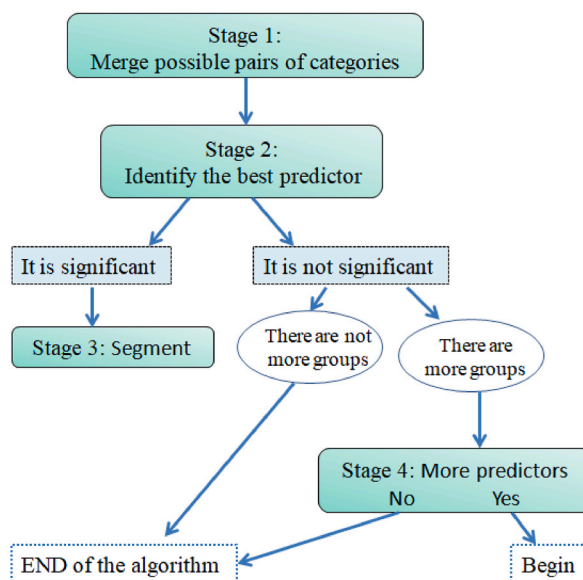


Fig. 2. Development of the CHAID algorithm.

CHAID can produce more than two categories at any level on the tree, whereas CART and QUEST only cater for a binary tree.

This study uses a data mining application of decision tree with CHAID algorithm with for identifying those predictors that are more informative about the sense of risks in the internet use.

#### 4. Results

The following are the social media most widely used by the young people taking part in this study: WhatsApp, (93.6%), YouTube (90.2%), Instagram (82.5%), and TikTok (65.5%). By contrast, high numbers reported making no use of platforms such as Facebook (89.9%), Telegram (89.5%), and Snapchat (74.1%).

In turn, 67.4% of the young people report that going online sometimes or often distracts them from their studies, while 55.2% have said that they sometimes or often feel that their use of the internet, social media or videogames is a waste of time. Finally, it should be noted that 89.1% of those surveyed declare that they have never suffered any online abuse, and 81.6% state that they have never been the victims of online bullying, as opposed to 15.6% that admit they have received online abuse.

The logistic regression analysis (Table 2) revealed the following factors linked to a sense of risk: gender, marks (academic record), Instagram, TikTok, Twitter, videogames, and pornography, distraction from studying, feeling isolated when not online, and using media serves no purpose. No association was found, however, with social media such as YouTube, WhatsApp, Spotify, Telegram, Snapchat, music, and betting.

The statistically significant variables in the logistic regression mode were used to locate the profiles experiencing a sense of risk online by a segmentation analysis based on the CHAID algorithm.

The average percentage of overall correct classification and the risk value of misclassification in the 10-fold cross validation was 35.4% (S.D. = 0.54) and values for cross validation were 38.6% (S.D. = 0.98). The training sample correctly classified 64.6% of young people, and the test-sample revealed that it correctly classified 61.4% of young people. This produced 13 nodes, of which eight are terminals (Fig. 3). The classification tree shows (node = 0) shows that 61% of the sample of young people experience a sense of risk. This sense is predicted by the variable “Feeling isolated when not online” ( $\chi^2 = 59.91$ ;  $p = 0.000$ ;  $df = 2$ ), noting that 68.3% of the young people that do not at all feel isolated do in fact perceive a sense of risk online compared to 31.7% that declare they do not (node

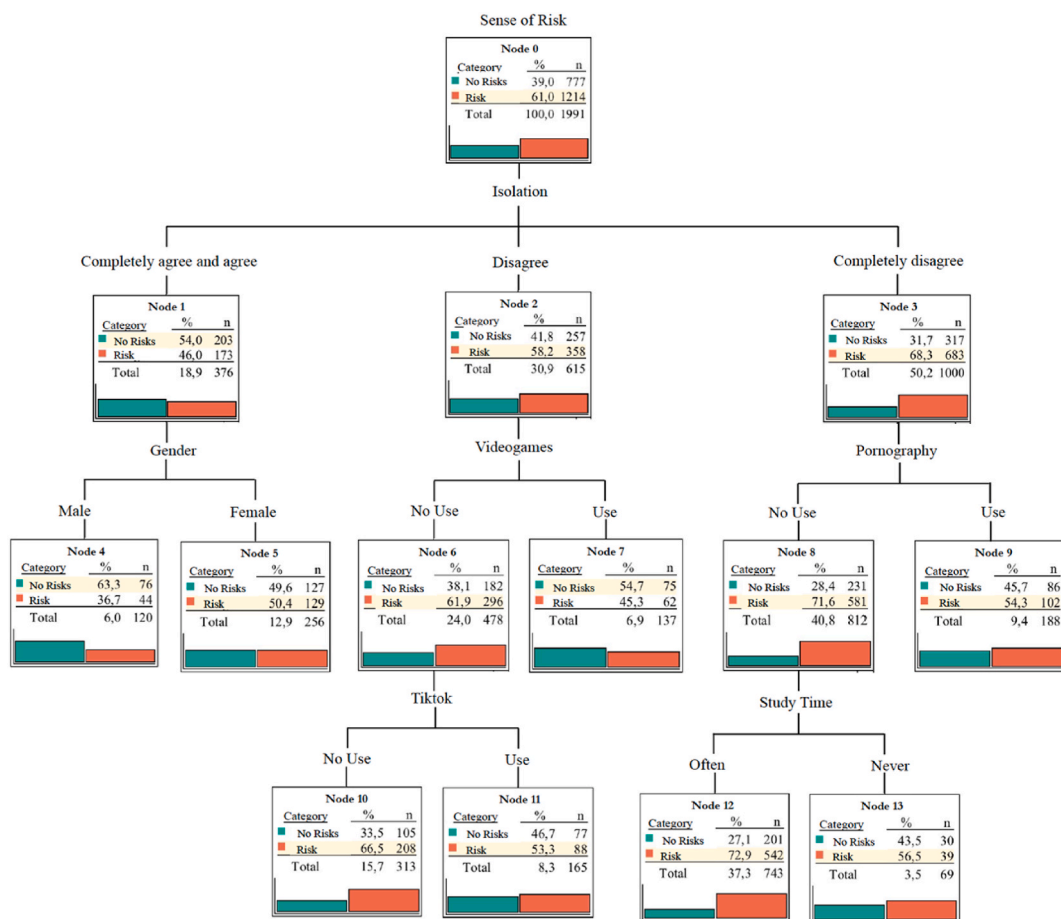


Fig. 3. Classification tree of the sense of online risk.



3). Within this segment, pornography is the second most predictive variable ( $\chi^2 = 21.093$ ;  $p = 0.000$ ;  $df = 1$ ). This means that the highest percentage in this group experiencing a sense of risk (71.6%) involves those individuals that either do not use or make only a limited use of the internet, and within this group the study control variable introduces a new significant differentiation ( $\chi^2 = 8.369$ ;  $p = 0.008$ ;  $df = 1$ ), with a greater sense of risk among those that have at some time or often been distracted from their studies. The highest percentage in this latter group is 72.9%.

An analysis of the segment comprising those young people that fully or fairly agree with feeling isolated, most of whom do not experience a sense of risk (54%), reveals that the differences are gender-based ( $\chi^2 = 6.195$ ;  $p = 0.013$ ;  $df = 1$ ), with males recording a higher percentage of those that do not experience a sense of risk (63.3%).

Finally, analysing the segment of young people that agree little with feeling isolated (58.2% perceive a sense of risk), the CHAID algorithm finds significant differences regarding the videogame variable ( $\chi^2 = 12.163$ ;  $p = 0.001$ ;  $df = 1$ ) and TikTok ( $\chi^2 = 7.888$ ;  $p = 0.010$ ;  $df = 1$ ) forming three terminal groups. The first group (node 10) consists of young people that agree little with feeling isolated and make limited or no use of videogames (61.9%) and TikTok. The second group (node 11) makes extensive use of TikTok (53.3%), little or no use of videogames, and agrees little with feeling isolated. The third and final group (node 7) is made up of young people that make extensive use of videogames and agree little with feeling isolated and do not experience a sense of risk (54.7%), compared to 45.3% that do indeed have this feeling (see Fig. 3).

## 5. Discussion

Based on the results forthcoming, we may posit that our study highlights a series of significant aspects that are consistent with prior studies conducted on this topic. Firstly, it has been noted that the most popular social media -scenarios of social interaction-among the young people taking part in this study are WhatsApp, YouTube, Instagram, and TikTok, whereas media such as Facebook are less popular, coinciding with the results recorded in other studies (Anderson & Jiang, 2018; Auxier & Anderson, 2021).

Secondly, this study has found that 15.6% of young people have been the victims of online bullying, which is slightly higher than in the study conducted by Garaigordobil and Larrain (2020), which reports a figure of 7.2%, and similar to the one conducted by Aboujaoude et al. (2015), in which the percentages of cyberbullying range between 20% and 40% of the young people taking part.

Thirdly, it should be stressed that young people's sense of risk linked to the internet, besides being predicted by the social isolation variable (*feeling isolated when not online*), was forecast by other variables that have been reported in prior studies as variables of risk associated with the internet. In particular we found risk variables such as study time, in which the internet is seen as a distraction (Plaza de la Hoz, 2021), gender, singling out differences in the use and risks associated with it depending on gender (Savoia et al., 2021), the consumption or not of pornography (Chen et al., 2021), the use of social media such as TikTok, and the use or not of videogames (Escario & Wilkinson, 2020; Masanet et al., 2021).

Fourthly, we find that the sense of risk appears to be predicted by a feeling of isolation when not online coinciding with similar results in prior studies, such as the ones by Álvarez and Moral (2020), Benchea (2021), and Marzilli et al. (2020). In our case, this variable was furthermore predicted by the participants' gender, with a highlight in our case being that those young people that identified themselves as male were those that reported the lowest experience of an internet-related sense of risk, coinciding with the results reported by Firat Sipal, Karakaya, and Hergul (2011). Moreover, other variables that predict the sense of risk were the use of videogames and TikTok, as well as the use of pornography and study time. These findings may be of special interest for those educational agents responsible for planning and implementing programmes and actions that involve the prevention and treatment of internet-related risks, as they allow identifying the profile of young people liable to fall into the dangers that online use entails.

Interestingly enough, our findings this time reveal that most of the young people that say they do not feel isolated when they are not online appear to perceive the risk associated with its use. Remaining with this same variable, we find gender-related differences, as the males in the group of young people that fully agree with feeling isolated if they are not online, males are the ones that mostly fail to perceive those risks. In turn, most of the young people that do not consume pornography, while also admitting that the internet has distracted them from their studies, also state that they perceive a sense of risk. The same occurs when we consider the group of young people that say they do not use videogames or TikTok; most of them say they experience a sense of risk regarding their online use.

In short, and adopting an education-related approach to our findings, it is of some concern that those young people that do not perceive this sense of risk associated with the use of the internet are precisely those that feel isolated when they are not online, or that those agreeing little with feeling isolated when not online make intensive use of videogames. In other words, precisely those young people that feel a greater need to go online are the ones that experience less the risks associated with online use. Moreover, our results enable us to confirm that, as stated in our study's theoretical underpinnings, we are dealing with a young cohort that lives online (Floridi, 2015; Heidari et al., 2020), and most of whom lack critical thinking and sense when they are and act online (Muñoz-Rodríguez, Dacosta, & Martín-Lucas, 2021; Prensky, 2001; Sánchez-Rojo, García del Dujo, Muñoz-Rodríguez, & Dacosta, 2022).

We therefore agree with the findings of prior studies (Torrijos Fincias, Serrate González, Martín Lucas, & Muñoz Rodríguez, 2021), as the relationship between youth and the internet necessarily involves educating young people in skills linked to consummate critical thinking that will enable them to convert risks into opportunities and proceed to confidently build their own identities. This involves a series of aspects that in pedagogical terms have their correlate in considering the approach to technology as an end in itself, for as this study has shown, technologies such as the internet have their own direct consequences regardless of the use we make of them.

## 6. Conclusions

Technologies such as the internet have now become the focal point where young people interact and socialise, ultimately constituting the scenarios in which adolescents live, communicate and spend a great deal of their time. Nevertheless, although this group often makes out –or wants to make out– that they have everything under control, the reality can sometimes be quite different. Whether it is because of the ease of going online, the freedom of movement from certain scenarios to others linked to its use, or the exposure to certain social relationships, the truth is that young people use and explore this technology at a frantic pace. According to [Gui et al. \(2023\)](#), this means that for as long as the internet remains part of young people's social, academic, working and family lives, there is a need to continue conducting studies to help us understand the influence that these kinds of technologies have on a crucial stage of development for building the identity of future generations.

Our research highlights two very significant aspects. Firstly, we present the profile of those young people that do not perceive the risk associated with online use, who are mostly boys that feel isolated when they are not online, as well as being consumers of videogames, pornography and social media such as TikTok. A second highlight is precisely the fact that those young people that most feel the need to go online are those that are not capable of perceiving the risks this technology entails, regardless of the use being made of it.

These results therefore reveal that, on the one hand, educational interventions addressing the risks associated with online use should adopt a positive attitude towards the use of technologies, such as the internet, as they are here to stay, with future generations already living online now and are set to continuing doing so in the future. On the other hand, it is therefore vitally important to undertake more research along these lines that will shed light on the development and influences technologies such as the internet are having, and indeed will continue to have, on young people's lives, as such technologies will be part of their process of adapting to a new context, a new understanding of space and time, made up of virtual communities and networks. A screen is not simply another device in a young person's hands, but instead the scenario, environment or context that caters for the inclusion of enculturation processes, as communication, interaction and action take place on and through it. It is therefore a cultural tool in young people's identity building; that is, a new way of accumulating social and cultural capital. We are not dealing with a technological issue, but instead an educational challenge that requires pedagogical approaches to implement a critical mediatic education that develops a responsible and unfettered identity, with critical and civic online autonomy.

Besides gathering results that provide us with a snapshot of young people in the 21st century, there is a need to simultaneously develop an educational mindset and praxis, as it is not only a matter of teaching young people to manage the use of their online time, but also to learn to understand the dangers associated with this technology. In educational terms, it would be more expedient to further generate critical thinking among young people and furnish them with strategies to enable them to forge a healthy and responsible relationship with technologies such as the internet. We therefore contend that we need to go beyond an education based on the restrictions of time and usages. We need to explore educational measures that show young people how technology can realign us as users, reboot our minds, adapting to new areas of interaction. Future generations need to be aware that regardless of how or why we use this technology, it entails a series of negative consequences that many young people are simply unaware of. In short, it requires developing young people's identities in response to today's technological (r)evolution, overcoming the mechanisms of digital literacy that have proven to be insufficient. We need to realise that in today's world, the boundary between online and offline is becoming increasingly blurred; we need to consider the risks a young person is exposed to in their online.

## 7. Limitations

Finally, we should like to mention this study's limitations. Firstly, by its very nature, our study purpose is subject to permanent and rapid changes, specific to today's social set-up and the swift development of technological applications. This means that our conclusions here are subject to an obvious risk of obsolescence in the medium run. Secondly, choosing schools in specific regions as our sampling points may lead to biases arising from the part played by the education authorities in the process of distributing the survey. Third, in this study a specific datamining method cannot be confirmed here because the cohort of individuals has not been chosen randomly.

### Credit author statement

Muñoz-Rodríguez, J.M.: Conceptualization, Funding acquisition, Project administration, Writing - original draft, Investigation. Patino Alonso, C.: Data curation, Formal analysis, Writing - original draft, Writing - review and editing, Software, Validation, Investigation. Pessoa, T.: Conceptualization, Writing - original draft, Martín-Lucas, J.: Conceptualization, Methodology, Writing - original draft, Writing - review and editing, Software, Visualization, Investigation.

### Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

### Data availability

Data will be made available on request.

## Acknowledgements

This work was supported by “CONNECT-ID. La identidad hiperconectada de la juventud y su percepción del tiempo en el ocio digital” [Hyperconnected identity of young people and their perception of time in digital pastimes], funded by Spain’s Ministry of Science, Innovation and Universities, grant number (Ref. PGC 2018-097884-B-I00), “NATEC-ID. Análisis de los procesos de (des-) conexión con la NATuraleza y con la Tecnología en la construcción de la IDentidad infantil” [Analysis of the processes of (dis-) connection with NATURE and TECHNOlogy when building a child’s IDentity] funded by Spain’s Ministry of Science and Innovation (PID 2021-122993NB-I00), and “Tecnología disruptiva como catalizadora de la transición ecológica desde la educación ambiental. estudio y diseño de soluciones tecnoeducativas desde NaturTEC Kids livinglab” funded by Spain’s Ministry of Science and Innovation (TED 2021-130300A-C22).

## References

- Aboujaoude, E., Savage, M. W., Starcevic, V., & Salame, W. O. (2015). Cyberbullying: Review of an old problem gone viral. *Journal of Adolescent Health, 57*(1), 10–18. <https://doi.org/10.1016/j.jadohealth.2015.04.011>
- Al-Furaih, S. A. A., & Al-Awidi, H. M. (2021). Fear of missing out (FoMO) among undergraduate students in relation to attention distraction and learning disengagement in lectures. *Education and Information Technologies, 26*, 2355–2373. <https://doi.org/10.1007/s10639-020-10361-7>
- Alkhasawneh, M. S., Kalthum Ngah, U., Tien Tay, L., Ashidi Mat Isa, N., & Subhi Al - Batah, M. (2014). Modeling and testing landslide hazard using decision tree. *Journal of Applied Mathematics, 2014*, 9. <https://doi.org/10.1155/2014/929768>. Arti.
- Álvarez, M., & Moral, M. V. (2020). Phubbing, uso problemático de teléfonos móviles y de redes sociales en adolescentes y déficits en autocontrol. *Health and Addictions, 20*(1), 113–125. <https://doi.org/10.21134/haaj.v20i1.487>
- Anderson, M., & Jiang, J. (2018). May 31). Teens, Social Media & Technology 2018. *Pew Research Center: Internet, Science & Tech*. Retrieved from <https://www.pewresearch.org/internet/2018/05/31/teens-social-media-technology-2018/>.
- Astleitner, H., Bains, A., & Hörmann, S. (2023). The effects of personality and social media experiences on mental health: Examining the mediating role of fear of missing out, ghosting and vaguebooking. *Computers in Human Behavior, 138*, Article 107436. <https://doi.org/10.1016/j.chb.2022.107436>
- Auxier, B., Anderson, M., & Pew. (2021). April 7). Social Media Use in 2021. *Pew Research Center: Internet, Science & Tech*. Retrieved from: <https://www.pewresearch.org/internet/2021/04/07/social-media-use-in-2021/>.
- Benchea, L. (2021). Analysis on the relationship between internet use and students academic engagement. *Journal of E-Learning and Higher Education. https://doi.org/10.5171/2021.237803*
- Bernal, A. (2021). Shaping identities in a hyperconnected world: Notes for the refutation of ‘pedagogical levity. In J. M. Muñoz-Rodríguez (Ed.), *Identity in a hyperconnected society. Risk and educative proposals* (pp. 3–14). Springer.
- Breiman, L., Friedman, J. H., Olshen, R. A., & Stone, C. J. (1984). *Classification and regression trees*. Wadsworth & Brooks; CERN Library.
- Brough, M., Literat, I., & Ikin, A. (2020). Good social media?": Underrepresented youth perspectives on the ethical and equitable design of social media platforms. *Social Media + Society, 6*(2), 1–11. <https://doi.org/10.1177/2056305120928488>
- Byrne, V. L. (2021). Blocking and Self-Silencing: Undergraduate Students’ Cyberbullying Victimization and Coping Strategies. *TechTrends, 65*, 164–173. <https://doi.org/10.1007/s11528-020-00560-x>.
- Calvete, E., Fernández-González, L., González-Cabrera, J., Machimbarrena, J. M., & Orue, I. (2020). Internet-risk classes of adolescents, dispositional mindfulness and health-related quality of life: A mediational model. *Cyberpsychology, Behavior, and Social Networking, 23*(8), 533–540. <https://doi.org/10.1089/cyber.2019.0705>
- Casaló, L. V., & Escario, J. J. (2019). Predictors of excessive internet use among adolescents in Spain: The relevance of the relationship between parents and their children. *Computers in Human Behavior, 92*, 344–351. <https://doi.org/10.1016/j.chb.2018.11.042>
- Çevik, Ö., Ata, R., & Çevik, M. (2021). Bullying and victimization among Turkish adolescents: The predictive role of problematic internet use, school burnout and parental monitoring. *Education and Information Technologies, 26*, 3203–3230. <https://doi.org/10.1007/s10639-020-10410-1>
- Chahal, H. (2013). ID3 modification and implementation in data mining. *International Journal of Computer Application, 80*(7), 16–23. <https://doi.org/10.5120/13873-1742>
- Chandra, S., Mondal, A., Ansary, K., & Saha, B. (2018). Social isolation in relation to internet usage and stream of study of under graduate students. *American Journal of Educational Research, 6*(4), 361–364. <https://doi.org/10.12691/education-6-4-10>
- Chen, J. K., Chang, C. W., Wang, Z., Wang, L. C., & Wei, H. S. (2021). Cyber deviance among adolescents in taiwan: Prevalence and correlates. *Children and Youth Services Review, 126*, 1–9. <https://doi.org/10.1016/j.childyouth.2021.106042>
- Escario, J. J., & Wilkinson, A. V. (2020). Exploring predictors of online gambling in a nationally representative sample of Spanish adolescents. *Computers in Human Behavior, 102*, 287–292. <https://doi.org/10.1016/j.chb.2019.09.002>
- Escofet, A., López, M., & Álvarez, G. (2014). Una mirada crítica sobre los nativos digitales: Análisis de los usos formales de TIC entre estudiantes universitarios. *REVISTA Q. Tecnología. Comunicación. Educación, 9*(17), 1–19.
- Evangelio, C., Rodríguez-González, P., Fernández-Río, J., & González-Villora, S. (2022). Cyberbullying in elementary and middle school students: A systematic review. *Computers & Education, 176*, Article 104356. <https://doi.org/10.1016/j.compedu.2021.104356>
- Firat Sipal, R., Karakaya, Y., & Hergul, N. (2011). I am online: What adolescents think about internet and internet use. *Procedia, Social and Behavioral Sciences, 30*, 2420–2426, 2011.
- Floridi, L. (2015). *The onlife manifesto: Being human in a hyperconnected era*. Springer.
- Garaigordobil, M., & Larrain, E. (2020). Acoso y ciberacoso en adolescentes LGTB: Prevalencia y efectos en la salud mental. *Comunicar, 62*, 79–90. <https://doi.org/10.3916/C62-2020-07>
- García del Dujo, Á., Vlieghe, J., Muñoz-Rodríguez, J. M., & Martín-Lucas, J. (2021). Thinking of (the Theory of) Education from the Technology of our Time. *Teoría de la Educación. Revista Interuniversitaria, 33*(2), 5–26. <https://doi.org/10.14201/teri.25432>.
- Gaspard, L. (2020). Australian high school students and their Internet use: Perceptions of opportunities versus ‘problematic situations. *Children Australia, 45*(1), 54–63. <https://doi.org/10.1017/cha.2020.2>
- Gómez Galán, J., Lázaro-Pérez, C., & Martínez-López, J. (2021). Exploratory study on video game addiction of college students in a pandemic scenario. *Journal of New Approaches in Educational Research, 10*(2), 330–346. <https://doi.org/10.7821/naer.2021.7.750>
- Gui, M., Gerosa, T., Argentin, G., & y Losi, L. (2023). Mobile media education as a tool to reduce problematic smartphone use: Results of a randomized impact evaluation. *Computers & Education, 194*, 104705. <https://doi.org/10.1016/j.compedu.2022.104705>
- Gür, D., & Kılıç Türel, Y. (2022). Parenting in the digital age: Attitudes, controls and limitations regarding children’s use of ICT. *Computers & Education, 183*, Article 104504. <https://doi.org/10.1016/j.compedu.2022.104504>
- Heidari, E., Salimi, G., & Mehrvarz, M. (2020). The influence of online social networks and online social capital on constructing a new graduate students’ professional identity. *Interactive Learning Environments. https://doi.org/10.1080/10494820.2020.1769682*
- Hodgson, N., Vlieghe, J., & Zamojski, P. (2020). Manifestations of the post-critical: From shared principles to new pedagogical paths. *Teoría De La Educación. Revista Interuniversitaria, 32*(2), 13–23. <https://doi.org/10.14201/teri.22576>.
- Hodkinson, P. (2017). Bedrooms and beyond: Youth, identity and privacy on social network sites. *New Media & Society, 19*(2), 272–288.
- Hu, J., & Yu, R. (2021). The effects of ICT-based social media on adolescents’ digital reading performance: A longitudinal study of pisa 2009, pisa 2012, pisa 2015 and pisa 2018. *Computers & Education, 175*, Article 104342. <https://doi.org/10.1016/j.compedu.2021.104342>

- INE (Spanish Office of Statistics). (2020). Población que usa Internet. Retrieved from [https://www.ine.es/ss/Satellite?L=es\\_ES&c=INESeccion\\_C&cid=1259925528782&p=1254735110672&pagename=ProductosYServicios/PYSLayout](https://www.ine.es/ss/Satellite?L=es_ES&c=INESeccion_C&cid=1259925528782&p=1254735110672&pagename=ProductosYServicios/PYSLayout).
- Kass, G. V. (1980). An exploratory technique for investigating large quantities of categorical data. *Journal of the Royal Statistical Society*, 29(2), 119–127. <https://doi.org/10.2307/2986296>
- Kesavaraj, G., & Sukumaran, S. (2013). A study on classification techniques in data mining, fourth international conference on computing, communications and networking technologies (pp. 1–7). ICCCNT. <https://doi.org/10.1109/ICCCNT.2013.6726842>, 2013.
- Kim, J. H. (2021). Factors associated with smartphone addiction tendency in Korean adolescents. *International Journal of Environmental Research and Public Health*, 18(21), Article 11668. <https://doi.org/10.3390/ijerph182111668>
- Kirschner, P., & De Bruyckere, P. (2017). The myths of the digital native and the multitasker. *Teaching and Teacher Education*, 67, 135–142.
- Liu, D., Wright, K. B., & Hu, B. (2018). A meta-analysis of social network site use and social support. *Computers & Education*, 127, 201–203. <https://doi.org/10.1016/j.compedu.2018.08.024>
- Li, J., Wu, Y., & Hesketh, T. (2023). Internet use and cyberbullying: Impacts on psychosocial and psychosomatic wellbeing among Chinese adolescents. *Computers in Human Behavior*, 138, Article 107461. <https://doi.org/10.1016/j.chb.2022.107461>
- Loh, W. Y., & Shih, Y. S. (1997). Split selection methods for classification trees. *Statistica Sinica*, 7, 815–840.
- Mace, R. (2020). Reframing the ordinary: Cyberspace and education. *Teoría de la Educación. Revista Interuniversitaria*, 32(2), 109–129. <https://doi.org/10.14201/teri.22473>.
- Malone, T. W., & Bernstein, M. S. (2015). *Handbook of collective intelligence*. MIT Press.
- Marzilli, E., Cerniglia, L., Ballarotto, G., & Cimino, S. (2020). Internet addiction among young adult university students: The complex interplay between family functioning, impulsivity, depression, and anxiety. *International Journal of Environmental Research and Public Health*, 17(21), 8231. <https://doi.org/10.3390/ijerph17218231>
- Masanet, M. J., Pires, F., & Gómez-Puertas, L. (2021). Riesgos de la brecha digital de género entre los y las adolescentes. *Profesional de la información*, 30(1), Article e300112. <https://doi.org/10.3145/epi.2021.ene.12>
- Memon, A., Sharma, S., Mohite, S., & Jain, S. (2018). The role of online social networking on deliberate self-harm and suicidality in adolescents: A systematized review of the literature. *Indian Journal of Psychiatry*, 60(4), 384–392.
- Muñoz-Rodríguez, J. M., Dacosta, A., & Martín-Lucas, J. (2021). Digital natives or digital castaways? Processes of constructing and reconstructing young people's digital identity and their educational implications. In J. M. Muñoz-Rodríguez (Ed.), *Identity in a hyperconnected society* (pp. 15–32). Switzerland: Springer.
- Muñoz-Rodríguez, J. M., Torrijos, P., Serrate, S., & Murciano, A. (2020). Entornos digitales, conectividad y educación. Percepción de la gestión del tiempo en la construcción de la identidad digital de la juventud. *Revista española de Pedagogía*, 78(277), 457–475. <https://doi.org/10.22550/REP78-3-2020-07>.
- Murciano Hueso, A., Gutiérrez Pérez, B. M., Martín Lucas, J., & Huete, A. (2022). Onlife youth: a study of young people's user profile and their online behaviour. *RELIEVE*, 28(2). <https://doi.org/10.30827/relieve.v28i2.26158>.
- Niemi, R. G., Portney, K. E., & King, D. (2008). Sampling young adults: The effects of survey mode and sampling method on inferences about political behavior. *Political Analysis*, 15(3), 286–323.
- Núñez, A., Álvarez-García, D., & Pérez-Fuentes, M. (2021). Ansiedad y autoestima en los perfiles de cibervictimización de los adolescentes. *Comunicar*, 67, 47–59. <https://doi.org/10.3916/C67-2021-04>
- Pangrazio, L. (2019). Technologically situated: The tacit rules of platform participation. *Journal of Youth Studies*, 22(10), 1308–1326.
- Plaza de la Hoz, J. (2021). Evolución del uso de las TIC por parte de los adolescentes en los últimos años: Luces y sombras. *Aloma, Revista de Psicología, Ciències de l'Educació i de l'Esport*, 39(1), 39–47. <https://doi.org/10.51698/aloma.2021.39.1.39-47>
- Prensky, M. (2001). Digital natives, digital immigrants. *On the Horizon. NCB University Press*, 9(5), 1–6.
- Resnik, F., & Bellmore, A. (2019). Connecting online and offline social skills to adolescents' peer victimization and psychological adjustment. *Journal of Research on Adolescence*, 48(2), 386–398. <https://doi.org/10.1007/s10964-018-0953-z>
- Sánchez-Rojo, A., García del Dujo, A., Muñoz-Rodríguez, J. M., & Dacosta, A. (2022). Grammars of “onlife” identities: Educational Re-significations. *Stud Philos Educ*, 41, 3–19. <https://doi.org/10.1007/s11217-021-09811-7>.
- Sánchez-Rojo, A., & Martín-Lucas, J. (2021). Education and TIC: Between means and ends. A post-critical reflection. *Educação & Sociedade*, 42, Article e239802. <https://doi.org/10.1590/ES.239802>.
- Savoia, E., Harriman, N. W., Su, M., Cote, T., & Shortland, N. (2021). Adolescents' exposure to online risks: Gender disparities and vulnerabilities related to online behaviors. *International Journal of Environmental Research and Public Health*, 18(11), 5786. <https://doi.org/10.3390/ijerph18115786>
- Selwyn, N., & Aagaard, J. (2020). Banning mobile phones from classrooms – an opportunity to advance understandings of technology addiction, distraction and cyberbullying. *British Journal of Educational Technology*, 52(1), 8–19. <https://doi.org/10.1111/bjet.12943>
- Soler, I. R., López-Sánchez, C., & Lacave, T. T. (2018). Online risk perception in young people and its effects on digital behaviour. *Comunicar*, 56, 71–79. <https://doi.org/10.3916/C56-2018-07>
- Telefonica Foundation. (2020). *Sociedad Digital en España*. Taurus.
- Thoilliez, B. (2020). When a teacher's love for the world gets rejected. A post-critical invitation to become an edifying educator. *On Education. Journal for Research and Debate*, 3(9).
- Torrijos Fincias, P., Serrate González, S., Martín Lucas, J., & Muñoz Rodríguez, J. M. (2021). Perception of Risk in the Use of Technologies and Social Media. Implications for Identity Building during Adolescence. *Education Sciences*, 11(523), 1–12. <https://doi.org/10.3390/educsci11090523>.
- Turkle, S. (2011). *Alone together. Why we expect more from technology and less from each other*. New York: Basic Books.
- Valdemoros, M. A., Sanz, E., & Ponce de León, A. (2017). Digital leisure and perceived family functioning in youth of upper secondary education. *Comunicar*, 50(25), 99–108.
- Vannucci, A., Simpson, E. J., Gagnon, S., & Ohannessian, C. M. (2020). Social media use and risky behaviors in adolescents: A meta-analysis. *Journal of Adolescence*, 79(1), 258–274.
- Vlieghe, J., & Zamojski, P. (2019). *Towards an ontology of teaching. Thing-Centred pedagogy, affirmation and love for the world*. Springer. <https://doi.org/10.1007/978-3-030-16003-6>
- Warden, C. A., Yi-Shun, W., Stanworth, J. O., & Chen, J. F. (2020). Millennials' technology readiness and self-efficacy in online classes. *Innovation in education and teaching international*, 59(2), 226–236. <https://doi.org/10.1080/14703297.2020.1798269>
- Yau, J. C., & Reich, S. M. (2019). It's just a lot of work: Adolescents' self-presentation norms and practices on Facebook and Instagram. *Journal of Research on Adolescence*, 29(1), 196–209. <https://doi.org/10.1111/jora.12376>
- Yildiz Durak, H., & Saritepeci, M. (2019). Modeling the effect of new media literacy levels and social media usage status on problematic internet usage behaviours among high school students. *Education and Information Technologies*, 24, 2205–2223. <https://doi.org/10.1007/s10639-019-09864-9>
- Zhou, M., & Zhang, X. (2019). Online social networking and subjective well-being: Mediating effects of envy and fatigue. *Computers & Education*, 140, Article 103598. <https://doi.org/10.1016/j.compedu.2019.103598>