

## Repositório ISCTE-IUL

---

Deposited in *Repositório ISCTE-IUL*:

2026-05-11

Deposited version:

Accepted Version

Peer-review status of attached file:

Peer-reviewed

Citation for published item:

Veiga Simão, A. M., Ferreira, P., Pereira, N., Oliveira, S., Paulino, P., Rosa, H....Trancoso, I. (2021). Prosociality in cyberspace: Developing emotion and behavioral regulation to decrease aggressive communication. *Cognitive Computation*. 13 (3), 736-750

Further information on publisher's website:

[10.1007/s12559-021-09852-7](https://doi.org/10.1007/s12559-021-09852-7)

Publisher's copyright statement:

This is the peer reviewed version of the following article: Veiga Simão, A. M., Ferreira, P., Pereira, N., Oliveira, S., Paulino, P., Rosa, H....Trancoso, I. (2021). Prosociality in cyberspace: Developing emotion and behavioral regulation to decrease aggressive communication. *Cognitive Computation*. 13 (3), 736-750, which has been published in final form at <https://dx.doi.org/10.1007/s12559-021-09852-7>. This article may be used for non-commercial purposes in accordance with the Publisher's Terms and Conditions for self-archiving.

---

### Use policy

Creative Commons CC BY 4.0

The full-text may be used and/or reproduced, and given to third parties in any format or medium, without prior permission or charge, for personal research or study, educational, or not-for-profit purposes provided that:

- a full bibliographic reference is made to the original source
- a link is made to the metadata record in the Repository
- the full-text is not changed in any way

The full-text must not be sold in any format or medium without the formal permission of the copyright holders.

---

# Prosociality in Cyberspace: Developing Emotion and Behavioral Regulation to Decrease Aggressive Communication

**Given name:** Ana Margarida      **Family name:** Veiga Simão <sup>a</sup>  
Mail: amsimao@psicologia.ulisboa.pt

**Given name:** Paula      **Family name:** Costa Ferreira <sup>a, b</sup>  
Mail: paula.costa.ferreira@gmail.com

**Given name:** Nádia      **Family name:** Pereira <sup>a</sup>  
Mail: nadia.msperreira@gmail.com

**Given name:** Sofia      **Family name:** Oliveira <sup>a</sup>  
Mail: sofia.oliveira94@outlook.com

**Given name:** Paula      **Family name:** Paulino <sup>a</sup>  
Mail: mailto:paula.paulino@ulusofona.pt

**Given name:** Hugo      **Family name:** Rosa <sup>b</sup>  
Mail: hugohrosa@gmail.com

**Given name:** Ricardo      **Family name:** Ribeiro <sup>b, c, \*</sup>  
Mail: ricardo.ribeiro@inesc-id.pt  
Telephone: +351 213100262  
Fax: +351 213145843

**Given name:** Luísa      **Family name:** Coheur <sup>b, d</sup>  
Mail: luisa.coheur@inesc-id.pt

**Given name:** João Paulo      **Family name:** Carvalho <sup>b, d</sup>  
Mail: joao.carvalho@inesc-id.pt

**Given name:** Isabel      **Family name:** Trancoso <sup>b, d</sup>  
Mail: isabel.trancoso@inesc-id.pt

<sup>a</sup> CICPSI, Faculdade de Psicologia, Universidade de Lisboa, Alameda da Universidade, 1649-013, Lisboa, Portugal

<sup>b</sup> INESC-ID Lisboa, Rua Alves Redol, 9, 1000-029, Lisboa, Portugal

<sup>c</sup> Iscte — Instituto Universitário de Lisboa, Avenida das Forças Armadas, 1649-026 Lisboa, Portugal

<sup>d</sup> Instituto Superior Técnico, Universidade de Lisboa, Av. Rovisco Pais, 1, 1049-001 Lisboa, Portugal

\* Corresponding author

**Conflict of Interest:** All authors declare that they have no conflict of interest.

# Abstract

## Introduction

Different forms of verbal aggression are often present in cyberbullying, which may impair executive function skills that enable the regulation of emotions and behavior. Emotion and behavioral regulation has been associated with better social adjustment and more positive interactions between peers. This study aimed to understand if fostering emotion and behavioral regulation strategies could decrease aggressive communication.

## Methods

A quasi-experimental longitudinal design, based on a Twitter client mobile application, with pre-posttest measures was used. For the application, we explore different machine learning approaches, including Computational Intelligence methods. Multilevel Linear Modeling and frequency analyses were performed. A convenience sample of 218 adolescents ( $M_{age} = 14.67$ ,  $SD = 0.84$ , 53% female) participated in the study.

## Results

Results suggest that a Twitter client mobile application intervention based on emotion and behavioral regulation strategies may help decrease adolescents' aggressive communication. Moreover, female and male participants who used the digital application tended to present distinct trajectories over time with regards to searching for information concerning prosocial behavior.

## Conclusions

These findings suggest that digital tools resorting to emotion and behavioral regulation strategies may be effective in reducing an aggressive communication style amongst adolescents and consequently, promote resource seeking to engage in prosociality. These results can be significant for the design of intervention programs against cyberbullying.

## Keywords

- aggressiveness
- behavioral regulation
- cyberbullying
- emotion in human-computer interaction
- prosociality

# 1. Introduction

Violence among peers has been a growing phenomenon worldwide amongst youth, making it a topic of priority. In a recent study, approximately 150 million adolescents have taken part in peer violence, such as sexual attacks, physical fights, bullying and cyberbullying [1]. Cyberbullying is considered an extension of bullying by some scholars and is often defined as intentional and repeated harmful behavior towards others through the use of electronic devices [2]. Although there are many benefits derived from the use of social media, namely improved subjective well-being from using social networks, individual differences seem to moderate positive or negative effects resulting from this use. Moreover, adolescents are often in contact with risk behavior while using social media such as cyberbullying [3] and quick maturational brain changes occur during adolescence, a time which is linked to risk-taking behavior and cognitive development [4].

Cyberbullying may have an impact on adolescents' mental health and well-being and, in turn, on their adjustment to school and academic performance [1]. Furthermore, this type of toxic experience may impair adolescents' development of executive function skills and self-regulatory behavior, which are important for them to manage threat and stress [5]. These executive function skills (i.e., working memory, inhibitory control and mental flexibility) are crucial for the development of emotion and behavioral regulation [6], which enable adolescents to plan and solve problems [7], which in turn, aid them in dealing with phenomena such as cyberbullying. Thus, it is essential to reduce adolescents' involvement in aggressive situations, such as cyberbullying and promote the self-regulation of emotion and behavior to foster healthy socialization experiences among adolescents [8].

Cyberbullying behavior may involve various forms of verbal aggression, such as threats and insults, which are used to inflict harm on others in virtual interactions [9]. Verbal aggression is related to an aggressive style of communication, which is an individual's personal characteristic that may be transferred to a communication act [10]. Affective-emotional processes have been known to be associated with aggressiveness and can inclusively be antecedents. Therefore, it is fundamental to understand how emotions can be regulated so that behavior may also be regulated to reduce aggression [11]. Effective emotion regulation is imperative for mental health, and difficulties in regulating emotions are related to injurious behavior [12]. Moreover, maladaptive emotion regulation, specifically suppressed anger, has been associated with aggressive behavior among youth [13]. In

fact, during adolescence, individuals' difficulty in regulating anger and sadness tends to be related with violence among peers, namely the use of physical and relational aggression [14].

Considering emotion and behavioral self-regulation processes enable individuals to self-direct their behavior toward more positive interactions with others [15], it is relevant to understand how self-regulation strategies may be used to reduce adolescents' aggressive behavior. Ultimately, this may prevent cyberbullying situations through less aggressive communication, and promote pro-social behavior among youth. In view of this, we propose to understand if a Twitter client mobile application that promotes emotion and behavioral regulation may have an impact on adolescents' style of communication when in online interactions. For this application, we explore different machine learning approaches, including Computational Intelligence methods [16, 17], an AI area that encompasses the theory, design, application and development of biologically and linguistically motivated computational paradigms (<https://cis.ieee.org/about/what-is-ci>). Artificial Intelligence (AI) approaches have the potential to make predictions across time and context, incorporating uncertainty in the analyses. This leads to informative predictions. Time is a key factor to understand how developmental processes occur [18]. Thus, when data is collected on numerous occasions for each individual, it constitutes a repeated measures design [19]. Accordingly, repeated measures data can be analyzed with multilevel modeling, where various levels (within and between subjects) may be examined, even with the incorporation of missing data, which is an advantage when examining ecologically valid data from real-world occurrences [18]. By comparison, biologically inspired AI for instance, is a promising avenue to explore complex non-linear phenomena with changing conditions [20]. Moreover, AI techniques, such as machine learning, are successful in analyzing biological systems and behavior, which are complex and hierarchical in nature and include multilevel data [21] — as is multilevel modeling [18]. Continuous data collection, as we present in this study, enables the analysis of behavior by both explained AI approaches [22] and multilevel modeling. Moreover, since the literature has mentioned that cyberbullying behavior seems to vary with regards to sex [23], we propose to better understand the type of impact the application may have on male and female adolescents. Therefore, in this research, we propose to answer the following questions:

- *Can there be change in an aggressive style of communication through the self-regulation of emotion and behavior?*
- *Are there any differences between male and female adolescents with regards to the application's usage?*

## 2. Cyberbullying: Communicating with Aggressiveness

The literature has linked aggressive behavior to a combination of deficits in assertiveness and a predominance of aggressiveness in communication. In cyberbullying, aggressive language is one of the main forms used to attack others by causing embarrassment, hurt, and psychological harm [9, 10]. The most frequent style of communication in cyberbullying is aggressiveness, which may constitute verbal attacks (e.g., on intelligence and physical appearance), insults and threats [23, 24].

Communication styles may be defined as cognitive processes that involve micro behavior which transmits literal meaning from one individual to another, and that include unique features which are produced in the act of communication [25, 26]. There are three key styles of communication which are associated with unique forms of verbal and nonverbal communication, namely assertiveness, passivity, and aggressiveness [27]. Accordingly, assertiveness is the expression of personal thoughts, feelings, opinions and needs in a direct, honest, and adequate manner. Passivity refers to not being able to express personal needs while denying personal rights. Lastly, aggressiveness is a style of communication which individuals use to claim personal needs and desires without respecting others. This paper focuses on this last style of communication, since it is predominant in incidents of cyberbullying [23, 24].

## 3. Fostering Prosociality through the Emotion and Behavioral Regulation of Aggressive Communication

Aggressiveness may be explained by the combination of impulsivity and the lack of mechanisms of emotion and behavioral self-regulation, reflected in the tendency to respond rapidly without thinking about behavior [28]. Furthermore, recent research has related difficulties in emotion regulation with cyberbullying behavior [29, 30]. Thus, using emotion and behavioral self-regulation strategies could be effective in helping adolescents regulate an aggressive style of communication, particularly when communicating with peers. Furthermore, reducing an aggressive style of communication could help adolescents cope with conflicts in a more pro-social way online [31] as well as in other interpersonal interactions (e.g., school, home). Ultimately, this could help reduce the incidence of cyberbullying amongst adolescents.

Self-regulation allows individuals to deal with the effects of external influences, and concurrently determines intentional action, thus, allowing to self-direct change [15]. Self-regulation processes derive from the level of

development in executive function skills, such as working memory, inhibitory control, and mental flexibility [6] and are part of individuals' personal agency, which refers to self-directedness - that is, the influence they have on their own functioning and their surrounding environment. This process involves intentional proactive plans of action and possible strategies (i.e., intentionality) which are triggered by executive function skills [7], to establish goals and anticipate outcomes which guide and motivate one's efforts (i.e., forethought), to execute action plans and self-regulate (i.e., self-reactiveness), and to self-examine one's own functioning (i.e., self-reflectiveness) [15].

Emotion regulation involves individuals' efforts to maintain, inhibit and enhance emotional experience and expression [32-34]. Emotions are related to cingulate activity in the limbic system, which is important for regulating affect [35], and can be considered as physiological, behavioral and experiential response tendencies that gradually emerge [36], involving factual cognitions (beliefs) and a continuous cycle of appraisals, responses and reappraisals that function to improve adaptive behavior [37]. Since individuals' evaluations depend on whether events are perceived as one's own responsibility, are probable, unexpected, desirable, and even morally good or bad [38], emotion regulation may be either deliberate (strategic), or automatic [39]. While deliberate emotion regulation is influenced by explicit goals, involving reflection, effort, and attentional resources [40], automatic emotion regulation involves implicit goals and a lack of reflective decision-making [41]. Accordingly, emotion regulation enables individuals to function effectively in their environment [32].

When faced with difficult emotion experiences for instance, those who can use deliberate emotion regulation adaptively through different emotion regulation strategies, such as emotional awareness and acceptance [42], are still able to engage in goal-directed behavior [43, 44]. While emotional awareness refers to the knowledge one has of one's own emotions [42], acceptance involves embracing undesired feelings without changing or eliminating them to mitigate emotional distress [45]. On the other hand, those who use maladaptive emotion regulation (i.e., over- or under-regulation) in difficult emotion experiences are unable to engage in goal-directed behavior and block the emotion experience from continuing [46]. This process can lead to negative consequences for individuals' psychological, physiological, social functioning and goal attainment, and may result in aggressive behavior [46]. In general, the literature seems to highlight the importance of emotion and behavioral regulation to promote positive social interactions and reduce aggressive behavior among adolescents. Aggressive online communication represents such behavior and is a component of cyberbullying.

Several studies have suggested that digital tools (e.g., applications, serious games) for intervention in cyberbullying should consider integrating self-regulation methods and include training of emotion-related

variables (e.g., empathy) to change negative behavior into prosociality [47]. One of these tools addresses bullying amongst adolescents by fostering empathy (e.g., FearNot!; [48]) which refers to individuals' ability to understand and feel others' emotions, and focuses on cognitive and affective dimensions [49]. Specifically, this research showed that children who experienced the intervention with those digital resources revealed more cognitive empathy towards others and more positive perceptions of their school environment. Other studies have focused on resources to prevent and intervene in cyberbullying through a social cognitive perspective (e.g., [15]). One of those studies refers to a serious digital game to intervene in cyberbullying (i.e., Friendly Attac), which revealed positive effects related to pro-social behavior, such as helping others, although no significant results were found concerning empathy [47]. Moreover, Reflective Interfaces is an example of a digital resource to address cyberbullying from a self-regulation approach, which involves the use of messages to encourage self-reflection in online behavior [50].

In view of the literature presented, we argue that fostering behavioral and emotion regulation strategies (e.g., emotional awareness and acceptance) may help adolescents decrease their aggressiveness when communicating online in difficult emotional experiences, such as cyberbullying incidents, and consequently, may lead to more pro-social behavior among adolescents. Moreover, following the objective of integrating self-regulation strategies in digital tools, this study proposes to understand if a Twitter client mobile application aiming to foster behavioral and emotion regulation could contribute to more positive social interactions among adolescents by reducing aggressive communication. In line with this, we hypothesized that:

*There can be a decrease in an aggressive style of communication through the use of a digital application integrating emotion and behavioral self-regulation strategies.*

In addition, sex seems to predict the type of cyberbullying behavior [51]. According to the literature, there is also a tendency for male individuals to be more engaged in direct forms of aggression (e.g., physical and verbal), whereas females tend to engage more in indirect types of aggression, such as social isolation and intentional exclusion from a group [52]. Consistently, indirect types of aggression are more frequent in cyberbullying and more common among females, as opposed to males [51]. Another study found that, unlike males, females tend to engage in cyberbullying in groups [23]. From a neurological perspective, studies on sex differences have revealed that several areas of functioning of the brain (i.e., white matter anisotropy) in male individuals seem to be related to lower impulsivity [53] and that maturation of medial frontal cortices in these individuals are associated with the development of rational decision-making and avoidance of risky choices [54]. These sex differences have implications in terms of specific vulnerabilities to maladaptive behavior,

considering their association with risk-taking tendencies [55]. In fact, considering the evidence presented, it seems that adolescent males may be less vulnerable than female adolescents to risk- and reward-related maladaptive behavior [55]. Therefore, we propose that the usage of the application may vary depending on whether adolescents are male or female. Hence, we argue that:

*There are differences between male and female adolescents with regards to the application's use.*

## 4. Method

### 4.1. Participants

A convenience sample of 218 9<sup>th</sup> graders ( $M_{\text{age}} = 14.67$ ,  $SD = 0.84$ , 53% female) from three public schools in Lisbon participated in this study. Adolescents' participation depended on their own volunteerism and parental consent. Participants were randomly distributed in four different conditions:

- Experimental condition 1 (EC1;  $n = 39$ ,  $M_{\text{age}} = 14.64$ ,  $SD = 0.87$ , 69% male);
- Experimental condition 2 (EC2;  $n = 64$ ,  $M_{\text{age}} = 14.83$ ,  $SD = 0.92$ , 64% female);
- Experimental condition 3 (EC3;  $n = 66$ ,  $M_{\text{age}} = 14.65$ ,  $SD = 0.87$ , 57% female);
- Control condition (CC;  $n = 49$ ,  $M_{\text{age}} = 14.53$ ,  $SD = 0.65$ , 51% male).

The imbalance concerning the size of the conditions resulted from practical constraints related to the school setting. Specifically, although participants were equally assigned to the different experimental conditions at the beginning of the investigation, a lack of parental authorization to use the application in the EC1, led to this imbalance. Ethically, since the participants from the different experimental conditions were in their natural groups (i.e., classes) and were all entitled to participate in the research project, they could not be moved to other conditions to ensure a balance in group size.

### 4.2. Instruments

#### 4.2.1. Measures

During the initial development of the measures used in the present investigation, the face and content validity were tested with three adolescents. The instruments were developed within two research projects<sup>1</sup>.

---

<sup>1</sup> *Cyberbullying: The regulation of behavior through language* [Foundation for Science and Technology (FCT),

*Aggressiveness in Interpersonal Communication* (AIC) is a 10-item task, based on Jakubowski and Lange's [27] theory of communication styles, specifically with regards to aggressiveness. It was assessed with the Item Response Theory (IRT) approach by computing Rasch analysis with the Winsteps program [56] because we wanted to measure its unidimensionality, including the level of difficulty of the tasks' items, as well as to understand participants' scores of aggressiveness in hypothetical situations. Participants were asked to respond to daily life situations of adolescents (e.g., 'A friend says to you [Your shirt is really horrible!]. How likely are you to respond? [What's that got to do with you?]). In doing so, they should position themselves in each situation and respond in the manner presented in terms of likelihood from 1 (not likely at all) to 5 (very likely). Winsteps enabled us to estimate the participants' scores on a one-dimensional logit scale and evaluate the properties of the AIC. We used Rasch polytomous methodology to examine the instrument and the participants' scores because we had polytomous data, rather than binary data, and wanted to avoid any effects of guessing due to the multiple-choice format of the questions in the task [57]. Specifically, we used the Partial Credit Model (PCM), an extension of the Rasch model for polytomous items [58]. The PCM for linear measures of observations of ordinal scales is  $\log(P_{nik}/P_{ni(k-1)})/\Theta_n - \beta_i t_{ki}$ , where  $P_{nik}$  is the probability that person  $n$  responds in category  $k$  when faced with item  $i$ . In accordance,  $P_{ni(k-1)}$  is the probability that the response is in category  $k-1$ ,  $\Theta_n$  is the ability of person  $n$ ,  $\beta_i$  is the difficulty of item  $i$ , and  $t_{ki}$  is the step calibration in the rating scale threshold (which is defined as the position equivalent to the equal probability of responses in adjacent categories  $k-1$  and  $k$  [59]). All items were assessed to understand whether they had excessive infit and outfit mean square residuals. None of the items showed infit/outfit higher than 1.5, as well as  $z$  statistic  $> 2.00$ , as suggested in the literature [60]. Item 2 was the easiest or the most reported item with a reported/difficulty level of  $-1.05$  log, whereas the most difficult or least reported was Item 8 with a reported/difficulty level of  $.96$  log. The distribution revealed a narrow range of difficulty ( $-1.05 < Di < .96$ ). We considered other reliability indicators from the Rasch measures for AIC including, Cronbach's alpha, Person Separation Reliability (PSR) and the Item Separation Reliability (ISR) [61]. AIC revealed a Cronbach's  $\alpha$  of  $.72$ , a PSR of  $.69$ , and an ISR of  $.99$ . These scores indicate good internal consistency reliability [62] even though the PSR revealed difficulty on the participants' behalf. After removing subjects with excessive infit/outfit, the PSR was  $.70$ , the ISR remained stable ( $.99$ ) and Cronbach's  $\alpha$  increased to  $.73$ .

---

PTDC/MHC/PED/3297/2014] and *The Bystander Effect* in Cyberbullying - taking responsibility and intervening through the regulation of behavior in adolescence (FCT, SFRH/BPD/110695/2015).

## 4.2.2. The ComViver Online Application

This Twitter client mobile application<sup>2</sup> was developed based on the principles of self-regulation to intervene in the phenomenon of cyberbullying. Twitter was chosen since it contains public content, thus, making it possible to extract data within a limit defined by Twitter itself, and because it was mainly used by adolescents (13-18 years). Initially, the application included an introductory communication practice exercise which informed users of the theme of the application. To fulfill its goal, the ComViver Online application provided opportunities for adolescents to engage in intentional forethought [15]. That is, the application included two integrated automatic detection systems (Support Vector Machines) which focused on classifying text regarding the presence of: i) aggressive language in isolated tweets, both when the user sent/received a tweet and when he/she browsed through his/her feed (i.e., Aggressiveness model); and ii) conflicts and/or attacks when tweets were part of an on-going public conversation between two or more users, that is, when users interacted with each other in specific contexts [63] (i.e., Conflicts/Attacks model).

To develop the ComViver Online application, we initially developed a tweet dataset in Portuguese (i.e., 40,000 tweets identified as containing potential aggressions from a set of 170 million initial tweets). We classified 5850 “aggressive” tweets from 40,000 tweets which were filtered by two researchers. This sample included 2,200 users. From the 170 million initial tweets, 58,470 tweets from these users were automatically classified as “non-aggressive”. This data set was used to train an automatic classifier concerning the detection of aggressive language. A second set of 1435 blocks of tweets without conflicts and/or attacks and 79 blocks containing conflicts and/or attacks was classified by five researchers to train the model of conflict and/or attack detection. Inter-rater reliability was .98.

Both predictive models were trained with samples of Portuguese tweets which were labeled by a total of seven independent educational psychology researchers, with experience in cyberbullying research, for the presence of aggressive language (i.e., single tweets) or conflicts/attacks (i.e., in case of an interaction). To identify aggressive language in online communication, the following criteria were used: a) the comment had to have the presence of offensive expressions/words; b) the content of the comment had to express an intention to harm

---

<sup>2</sup> Developed within the project *Cyberbullying: the regulation of behavior through language* (FCT, PTDC/MHCPED/3297/2014).

others (intentionality); and c) the comment had to be directed towards a peer/group of peers. The tweets had to meet all three criteria for being classified as aggressive language.

The evaluation of the predictive models showed that the Aggressiveness model was 85% precise and 93% sensitive (recall), while the Conflicts/Attacks model was 93% precise and 84% sensitive. Every time each of the models detected tweets and/or interactions containing aggressive language and/or a conflict/attack, the user/s involved was/were identified and a set of technical features were mobilized, encouraging user/s to engage in forethought to self-regulate emotions and behavior, as he/she communicated with others online. Therefore, both models work as an intervention tool, more specifically as self-regulatory strategies which foster opportunities for adolescents to engage in forethought to regulate their behavior while they communicate with others.

Although these models constitute an improvement in the field of automatic detection of cyberbullying, they do not yet detect cyberbullying. Rather, they identify the presence of aggressive language and conflicts and/or attacks online. Thus, improvement needs to be done by focusing on events that go beyond a one-to-one relationship and also which capture the context of interpersonal online interactions [63]. Only then would it be possible to approximate models of automatic detection to actual cyberbullying events.

The automatic detection systems combined with the technical features, enabled the application to alert the sender prior to posting a message, that it may contain aggressive content and whether he/she wanted to send it, triggering a chain of options to promote forethought on the sender's behalf (see Fig. 1).

These features also masked posts that were potentially aggressive and alerted the receiver of this content, as well as questioned the latter whether he/she wanted to see the post, activating a chain of options to promote forethought and foster emotion regulation on the receiver's behalf (see Fig. 2). Prior to either posting or receiving a potentially aggressive message, this initial process provided users with a moment to self-reflect on the development of a plan of action and strategies to use, as well as anticipate an outcome so that personal goals may be achieved through pro-social behavior online, as opposed to aggressive behavior [15].

Moreover, the application provided a set of psychoeducational resources aiming to promote different self-regulation processes since self-regulation may be developed through dynamic cyclical phases and different self-regulation strategies [64]: Inspire Yourself, Express Yourself, Did you know, Know more about, True stories, and Videos. For every aggressive tweet detected, the application generated a psychoeducational resource as a fake tweet in the users' feed (i.e., both sender and receiver), to ensure that the users were exposed to the

resources in the same proportion they see aggressive messages. Additionally, users could consult all resources in the application menu autonomously.

For users to develop their self-reactiveness (i.e., executing action plans and self-regulating behavior; [15, 64]), Inspire Yourself (see Fig. 3) included a combination of images with quotes which intended to guide users toward pro-social behavior while fostering emotional awareness and acceptance (i.e., strategies of emotion regulation) which in turn, could buffer cyberbullying behavior [30]. Additionally, the process of self-reactiveness was promoted through Express Yourself (see Fig. 4), which was an interactive resource that included communication quizzes to prevent aggressive communication. This resource enabled users to view hypothetical scenarios to which they responded, and each response reflected a specific communication style (i.e., assertiveness, passivity or aggressiveness; [27]). Therefore, choosing more aggressive responses reflected aggressiveness in communication which seems to be related to antisocial and/or aggressive behavior and a lack of assertiveness [65, 66]. Since this resource was interactive, meaning that a feedback was provided to the users according with the responses they selected, it allowed them to self-reflect on their answers, to re-think their action plan and, consequently, to self-regulate their behavior in interpersonal interactions with others. Therefore, this resource also fostered self-reflectiveness [15], that is, the users' self-examination of their own functioning through specific feedback. Specifically, the feedback offered to the users on how they responded to the different scenarios either reinforced an assertive communication style or encouraged them to reflect on their choice and to retry differently in case their response reflected aggressiveness or passivity. Moreover, if the user was involved in a conflict/attack (as either sender or receiver) in the previous 48 hours, the Express Yourself resource doubled the likelihood of it appearing in the feed, as a way to prevent aggressive communication.

Seeking information and social assistance are key self-regulatory strategies in the resolution of cyberbullying situations, as previous research has indicated [67, 68]. Thus, the application also provided users with opportunities to seek information by clicking an option entitled "Information about cyberbullying" (see Table 1 for description of resources), as well as social assistance by clicking either on the option "Call an emergency contact" (e.g., family member/close friend) and/or "Call the Portuguese Association for Victim Support" (see Fig. 2).

During the development of the application, a face validation was conducted involving 16 9<sup>th</sup> graders ( $M_{age} = 14.50$ ,  $SD = 0.63$ , 63% male) from a public school in Lisbon. These participants tested an initial prototype of the application and adjustments were made according to their experience, leading to an improved version.

### 4.2.3. Procedures

We were granted authorization to conduct this study by the Ministry of Education of Portugal, the Portuguese National Commission of Data Protection, the Deontology Committee of the Faculty of Psychology of the University of Lisbon, the schools' boards of directors, the teachers, the parents, and the adolescent participants. The task on aggressiveness (i.e., AIC) was administered to adolescents in a classroom context with computers with Internet access in their own schools by the researchers of this study. We informed all students that they could have psychological support (i.e., with a psychologist) if they needed. Furthermore, we informed all students that they could quit at any time they wished to.

The intervention with the Twitter client mobile ComViver Online application had a duration of ten weeks and involved the participation of 9<sup>th</sup> graders who were randomly assigned to four conditions (i.e., EC1, EC2, EC3 and CC). In the first three sessions, all students were informed about the project, performed pretests and wrote down their expectations (except for the CC) with regards to the training with the application. All participants without a Twitter account were invited to create one at the beginning of the research, except for the CC participants, who did not use the resources or the application. In session 4, students from EC1 installed the application and learned how to manage it. From this session onward, students were able to use the application as they wished. To monitor students' application use, we set up sessions 5, 6, 7 and 8 to answer any questions or doubts students may have regarding the use of the application. These four sessions were conducted by two educational psychologists and a computer science engineer with expertise in cyberbullying to clarify doubts that could occur and ensure the application was running correctly. In sessions 9 and 10, all students performed a post-test and a general evaluation of the intervention. These sessions were on a weekly basis. Since students in the EC1 had the application installed on their cellphones, we were unable to prevent them from using it if they wished to. Our main goal with the intervention, however, was not to promote a more or less regular use of the social network where our application is embedded, but to assure that when using it, participants had access to a number of digital self-regulation resources that promoted reflection to address cyberbullying.

As for the alternative interventions (i.e., EC2 and EC3), the same educational psychologists conducted weekly sessions allowing the students to consult some of the psychoeducational resources that were embedded in the application, on paper. Both conditions had access to the psychoeducational resources that enabled users to seek information (i.e., Did you know, Know more about, True stories and Videos). However, the two conditions were distinguished based on whether they worked with "Inspire yourself" (EC3) or "Express yourself" (EC2). In each

session, students for each condition were organized in small workgroups that were randomly assigned at the beginning of each session so that they experienced different interpersonal relationships. The participants in the CC did not have contact with any of the resources or the application. With regards to the participants in the alternative intervention and control condition, it is not possible to compare app usage, since they did not have access to the app. We controlled the usage time of all participants in weekly sessions to make it equivalent to the monitoring sessions of the EC1.

#### 4.2.4. Data Analysis

Estimated marginal means concerning an aggressive communication style were calculated from a database which was previously treated for missing values by transforming raw data into *Z* scores and by removing outliers. To control pretest differences for an aggressive style of communication between the four conditions considering sex, age and class, we computed analyses of covariance (ANCOVA) with the pretest value as covariate with IBM's SPSS, 23.0. Levene's test and normality tests were computed to verify assumptions. A Levene test revealed that the assumption of homogeneity of variance was met ( $F = .90, p = .44$ ). Although the non-normal distribution of the variables is commonly found when working with psychological variables and psychometric measures [69], we followed the Central Limit Theorem, which states that given certain conditions (i.e., sample sizes  $> 30$ ), the distribution will tend to be normal, even if the sample is non-normally distributed [70]. The ANCOVA analysis enabled us to compare mean differences between groups after the intervention. To get a better understanding of how the application was used, we also computed Multilevel Linear Modeling (IBM, SPSS, 22.0) with logfile longitudinal aggregated data of the 39 students in EC1. We adopted this statistical procedure because we wanted to understand the potentially different temporal trajectories of the male and female participants' use of the application. For longitudinal analyses purposes, we used the logfile data from sessions 5, 6, 7 and 8. We opted to examine differences between male and female participants, since evidence has shown that the latter may engage in different types of online aggression [23] that adolescent males may be less vulnerable than female adolescents to risk- and reward-related maladaptive behavior [55].

A sample size of 156 session entries (i.e., four session entries per student measured on four occasions) was used for application use at level 1 and of 39 students at level 2. We computed Maximum Likelihood for estimation for all analyses, which offers asymptotically unbiased estimates [71] and introduced the variables in SPSS in three steps (i.e., an intercept-only model, a model with a linear trend and a model with differences in development between the two groups – male and female) to test the interaction effects. A scaled identity

covariance structure for the repeated measures effect and a variance components covariance structure for the intercept random effect were also used to examine the amount of variance in the outcome within and between individuals. The scaled identity covariance structure assumes that there is a constant variance across occasions with no correlation between components and has one estimated parameter [18]. We assessed the improvement of each model over the previous one with the corresponding likelihood ratios. Additionally, we computed percentages of the time adolescents dedicated to each specific resource.

## 5. Results

There was a significant difference in mean concerning an aggressive style of communication [ $(3,213) = 2.783, p < 0.05, \eta_p^2 = .04$ ] between conditions. In the ANCOVA analysis, the estimated mean for conditions regarding an aggressive style of communication in the pretest was 2.09. Post hoc tests showed there was a significant difference between the EC1 and EC2 ( $p < 0.05$ ). The estimated marginal means revealed that the group which revealed less of an aggressive style of communication at the end of the ComViver Online application, was the EC1 ( $M = 1.97$ ) compared to EC2, EC3 and CC ( $M = 2.27, M = 2.22, M = 2.14$ , respectively). Fig. 5 presents the results of the ANCOVA analysis to examine the effectiveness of the ComViver Online application with regards to an aggressive style of communication by comparing means between groups, before and after the intervention.

Results from the Multilevel Linear Modeling revealed that the male and female participants' trajectories were significantly different as sessions progressed. In fact, males used the self-regulation strategy of seeking for information regarding pro-social behavior in the resources available in the platform more than the females throughout time. Figure 6 shows both male and female participants' aggregated trajectories with regards to seeking information on prosocial behavior throughout time (in the different sessions). While the male participants' temporal trajectories evolve in an ascending manner throughout time, the female participants' temporal trajectories evolve in a descending manner throughout the different sessions.

At level 1, the variance corresponds to the variability in the average adolescents' time dedicated to seeking information with resources to guide them towards pro-social behavior estimates around their own growth trajectory [72]. The estimates of variance for levels 1 and 2 ( $Z_w = 7.64, p < .001; Z_w = 2.32, p < .001$ , respectively), suggest that there was enough variation in intercepts across adolescents. The proportion of variance (ICC) was computed with a one-tailed test for variances, which revealed a variance between

individuals of 23% and 77% of variance within individuals for adolescents' time dedicated to seeking information. Therefore, we concluded that there was variance within and between individuals over time. The intercept-only model, which included only the intercept, was compared with the intercept + time model (see Table 2; notice that Level 1 means and standard deviations are reported according to the time variable (from 0 to 3); standard errors are in brackets;  $\dagger p < 0.10$ ,  $* p < 0.05$ ,  $** p < 0.01$ ). The intercept + time model revealed a significant improvement over the intercept-only model:  $deviance = 3.51$ ,  $df = 1$ ,  $p < .01$ . In this second model, the intercept corresponds to adolescents' time dedicated to seeking information at the beginning of the study. A linear time variable was significant in explaining the growth of adolescents' time dedicated to seeking information. The model with the predictor variables and the interaction between them presented a significant improvement over the intercept + time model:  $deviance = 4.60$ ,  $df = 2$ ,  $p < .05$ . Specifically, males tended to search for more resources throughout time than females (an average increment of 15.62 for each unit of evaluated time). Table 2 shows the model fit information (likelihood ratios) and estimates for the fixed and random effects of all models.

To complement these results, we present the percentages pertaining to the time adolescents dedicated to specific resources to guide them towards pro-social behavior (see Table 3). There is variability regarding the time dedicated to different resources among female and male adolescents. In particular, the resource "Inspire Yourself" was sought twice as much by males than females.

## 6. Discussion

Results supported our hypothesis, indicating that a decrease in aggressive communication may have been promoted through the use of self-regulation strategies integrated in the Twitter client mobile application, in favor of the EC1 when compared to the other three conditions (i.e., EC2, EC3 and CC), although differences were only statistically significant between EC1 and EC2. These results complement previous studies that found that individuals' pro-social skills, self-efficacy and intentions to engage in positive bystander behavior in cyberbullying situations improved after experiencing a serious game intervention [47]. Accordingly, these previous studies suggested that research investigate negative online behavior and cyberbullying involvement, as we aimed to do with aggressive communication. Aggressive communication in the EC1 decreased with respect to EC2, EC3 and CC, although, as previously mentioned, the only statistically significant difference was between the EC1 and EC2. These results suggest that the application may have contributed to help adolescents self-regulate their aggressive communication through the set and combination of self-regulation strategies which

are integrated in this digital tool. This result concerning the decrease in the EC1 participants' aggressive communication adds to other findings that indicated that children who experienced an intervention with digital resources that promoted reflection to address bullying, revealed more cognitive empathy towards others and more positive perceptions of their school environment [48].

In addition, there were specificities between conditions which need to be considered to better understand these results. Firstly, the EC1 experienced more self-regulatory strategies than the other two conditions, which were involved in alternative interventions (i.e., EC2 and EC3). The application integrated models of automatic detection of cyberbullying combined with other self-regulatory features (i.e., forethought, seeking information and social assistance, reflecting on pro-social behavior and emotion regulation, and self-reflectiveness), as previous research has recommended [50].

In contrast, although EC2 experienced the same four resources which promoted seeking information and the "Express yourself" resource (i.e., all on paper), students in this condition did not receive any feedback regarding their choices. Therefore, we believe that the significant difference found between the EC2 and the EC1 was due to the fact that, even though this particular resource on paper (i.e., "Express yourself") was a self-regulatory strategy which fostered the process of self-reactiveness (i.e., executing action plans and self-regulating behavior; [15]), it did not provide any feedback to the users. Thus, it is possible that this resource on paper was not able to promote a self-reflection concerning users' responses and therefore, it did not promote self-reflectiveness (i.e., self-examination of one's own functioning; [15]), as it probably did in the EC1. In fact, it could have contributed to increase adolescents' aggressiveness, since they were exposed to aggressive choices without reflecting on communication.

Considering EC3, adolescents also experienced the four resources that fostered seeking information but were only in contact with the "Inspire yourself" resource on paper, as opposed to the "Express Yourself" resource. Thus, we believe that this resource may have contributed to a lower difference between EC3 and EC1 (i.e., not statistically significant). This may have had an impact on this condition participants' aggressive communication, as this resource aimed to promote self-reflection and emotion regulation. This finding complements previous research that provided evidence that reflection (i.e., reflective message) may prompt a decrease in individuals' intentions to harm others online in different experimental conditions [47]. Accordingly, while previous research focused on parents as audience, disapproval from others and potential harm for the receiver of the message in the different experimental conditions, we focused on how varied self-regulation resources (i.e., experimental conditions) may cause a change in aggressive communication.

Lastly, the less expected result pertained to the CC, which did not experience any of the resources or the application and revealed the lower difference in relation to the EC1 (i.e., not statistically significant), in comparison with EC2 and EC3. A possible explanation for this result may refer to uncontrolled differences between conditions (e.g., participation in other curricular or extracurricular activities impacting personal characteristics), since this group of adolescents did not participate in the monitoring sessions and, thus, was not regularly followed by the research team.

With regards to sex, this study revealed that males used the self-regulation strategy of seeking for information regarding pro-social behavior in the resources available in the platform more throughout time than females throughout the sessions. Male participants may have sought more information regarding pro-social behavior because they tend to engage in direct forms of aggression (e.g., physical and verbal) more than females, and the application focused more on training direct forms of aggression, (i.e., aggressive communication) as opposed to indirect forms of aggression, such as exclusion [52]. Another interpretation is that male adolescents seem to show lower impulsivity [53], increased rational decision-making and avoidance of risky choices when compared to females [54], and therefore, they may have searched for more information in the application over time as a problem-solving strategy. We also provided information regarding the specific resources males and females dedicated their time to. These varied between sex and between resource. These results hold prospects for developing programs that are tailored to adolescents, considering their individual differences, such as sex. The implications of these results pertain to how personalizing programs to meet adolescents' need may fall on considering male and female vulnerabilities to maladaptive behavior [55]. This personalization aspect of designing applications is fundamental because learning prosocial behavior should occur in a meaningful way. Since cyberbullying behavior is manifested and dealt with differently, depending on individuals' sex [51], it is crucial that digital educational applications should be designed according to individuals' needs. The present study provides an important contribution in respect to this.

These results contribute to the overall finding which indicate that the different and combined behavioral and emotion regulation strategies integrated in the Twitter client mobile ComViver Online application seem to have a potential to change an aggressive style of communication, by decreasing it.

## 7. General Conclusion

Peer violence such as cyberbullying, has been increasing worldwide, presenting a negative impact on youth's psychological adjustment and their relation with school [1, 3]. Since these experiences can have a negative impact on adolescents' development of executive function skills which are important for the development of emotion and behavioral regulation [6], it is essential to minimize adolescents' aggressiveness and consequent involvement in cyberbullying to foster healthy socialization experiences in adolescents [8]. Accordingly, difficulties in emotion regulation have been related with adolescents' aggressive behavior and peer violence [13, 14] and their predisposition to search for strategies to solve-problems may have been due to individual characteristics, such as sex [55]. From our results, it seems that emotion regulation along with the self-regulation of behavior help decrease an aggressive style of communication [11]. In this regard, using self-regulation strategies focusing on the promotion of pro-social behavior to intervene in this type of aggressive episodes may help reduce its incidence among youth.

Results from our study revealed the relevance of an emotion and behavioral self-regulated approach as an opportunity to prevent and intervene in cyberbullying in a pro-social manner (e.g., [15, 64]). Specifically, our findings suggest that if diverse and combined self-regulation features (e.g., forethought, seeking information and social assistance, reflecting on pro-social behavior and emotion regulation, and self-reflectiveness) are integrated in digital applications, they may reduce aggressive online interactions amongst adolescents and act as protective factors. Hence, Social Network Systems could integrate these features to generate safer environments for users. In accordance, emotion and behavioral regulation have the potential to help adolescents regulate how they communicate with others in both online and offline interpersonal relationships and social interactions. In accordance, an aggressive style of communication may be reduced among adolescents, not only in the context of cyberbullying intervention, but also within a broader scope of promoting pro-social behavior, emotional well-being and positive relationships among peers through self-regulation.

This study presents some limitations that should be addressed. Firstly, the size imbalance between the groups, which derived from particular features of the educational setting, already mentioned in Section 4, is an important limitation. Additionally, rather than including only cross-sectional data and pre-posttests, it would be interesting for future research to study these variables through longitudinal designs with follow-up measures focusing on individual characteristics, as well as on the effects of extracurricular activities in schools, such as the ones presented in this study. Also, since this study lasted eight weeks, it would be important to replicate it

for a longer period of time, since it may increase its efficacy [73], using repeated measures to monitor the possible changes in the variables assessed [74], namely, self-regulation and aggressive communication style. Another limitation of this study could be that even though psychologists had access to the students' tweets, as they monitored their use of the application, they did not label these tweets, as this was not an objective of this study. As such, unsupervised and graph algorithms could be considered in future studies to provide a deeper analysis of the tweets' specific content. That is, unsupervised learning methods could detect data patterns that could shed new light on the dynamics of verbal communication in cyberbullying contexts, since it does not require labeled data. With this method, groups of profiles could be created through clustering in a way that those within the same cluster would be similar to each other, and dissimilar to those in other clusters. This grouping could be used to identify potential cyberbullying victims, observers and cyberbullies, considering individual variables, such as sex and age. Moreover, future studies (follow-up) with the application could provide a precision/recall score of the models on the deployed application. Additionally, it would also be relevant for future research to provide a better understanding of the relation between automatic cyberbullying detection systems embedded in the ComViver Online application and specific emotion and behavioral regulation strategies. In fact, future directions with this Twitter client could include classifying usage trajectories in cyberbullying contexts throughout time with multilingual Computational Intelligence techniques that would enable a detailed analysis between Twitter users of different cultural backgrounds in different countries – thus creating an extensive hierarchic data structure.

To conclude, this study proposes the use of digital tools that foster emotion and behavioral regulation to effectively reduce adolescents' aggressive communication. Moreover, social media should be regarded not only as a risk context where adolescents are constantly exposed to potentially harmful situations, but also as an opportune mean to promote prosociality among them. Ultimately, this type of approach may contribute to diminishing the incidence of cyberbullying.

## 8. Acknowledgements

This work was supported by the Foundation for Science and Technology (FCT) of the Science and Education Ministry of Portugal (PTDC/MHC/PED/3297/2014; SFRH/BPD/110695/2015), along with the Research Center for Psychological Science (CICPSI; UID/PSI/4527/2016), and in collaboration with INESC-ID (UIDB/50021/2020; SFRH/BSAB/136312/2018).

## 9. Compliance with Ethical Standards

**Funding:** This study was funded by The Portuguese Foundation for Science and Technology (PTDC/MHC/PED/3297/2014; SFRH/BPD/110695/2015) of the Science and Education Ministry of Portugal, along with the Research Center for Psychological Science (CICPSI; UID/PSI/4527/2016), and in collaboration with INESC-ID (UIDB/50021/2020; SFRH/BSAB/136312/2018)

**Ethical approval:** All procedures performed in our study involving human participants were in accordance with the ethical standards of the institutional and/or national research committee (Portuguese Ministry of Education, National Committee of Data Protection, Schools' Board of Directors) and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

**Informed consent:** Informed consent was obtained from all individual participants included in the study, including their legal guardians.

## 10. References

- [1] UNICEF - United Nations Children's Fund. A Familiar Face: Violence in the lives of children and adolescents. UNICEF, New York, USA; 2017.
- [2] Patchin JW, Hinduja S. Bullies move beyond the schoolyard: A preliminary look at cyberbullying. *Youth Violence and Juvenile Justice* 2006;4(2):148–69.
- [3] UNICEF - United Nations Children's Fund. The State of the World's Children 2017: Children in a Digital World. UNICEF, New York, USA; 2017.
- [4] Weinstein A, Dannon P. Is Impulsivity a Male Trait Rather than Female Trait? Exploring the Sex Difference in Impulsivity. *Curr Behav Neurosci Rep* 2015;2:9–14.
- [5] Evans GW, Wachs, TD, eds. Chaos and its influence on children's development: An ecological perspective. Washington, DC: American Psychological Association; 2010.
- [6] Blair C, Diamond A. Biological processes in prevention and intervention: the promotion of self-regulation as a means of preventing school failure. *Development and Psychopathology* 2008;20(3):899-911.

- [7] Berthelsen D, Hayes N, White SLJ, Williams KE. Executive Function in Adolescence: Associations with Child and Family Risk Factors and Self-Regulation in Early Childhood. *Frontiers in Psychology* 2017;8:903.
- [8] Rueda MR, Posner MI, Rothbart MK. The Development of Executive Attention: Contributions to the Emergence of Self-Regulation. *Developmental Neuropsychology* 2005;28(2):573-94.
- [9] Veiga Simão AM, Ferreira P, Francisco SM, Paulino P, de Souza SB. Cyberbullying: Shaping the use of verbal aggression through normative moral beliefs and self-efficacy. *New Media & Society* 2018;20(12):4787–806.
- [10] Lin Y, Durbin JM, Rancer AS. Perceived instructor argumentativeness, verbal aggressiveness, and classroom communication climate in relation to student state motivation and math anxiety. *Communication Education* 2017;66(3):330–49.
- [11] Baumeister RF, Bushman BJ. Emotions and Aggressiveness. In: *International Handbook of Violence Research*, Springer Netherlands, Dordrecht; 2003. pp. 479–93. ISBN 978-0-306-48039-3.
- [12] Kun B, Demetrovics Z. Emotional intelligence and addictions: A systematic review. *Substance Use & Misuse* 2010;45(7-8):1131–60.
- [13] Norström T, Pape H. Alcohol, suppressed anger and violence. *Addiction* 2010;105(9):1580–6.
- [14] Sullivan TN, Helms SW, Kliewer W, Goodman KL. Associations between sadness and anger regulation coping, emotional expression, and physical and relational aggression among urban adolescents. *Social Development* 2010;19(1):30–51.
- [15] Bandura A. Toward a psychology of human agency. *Perspectives on Psychological Science* 2006;1(2):164–80.
- [16] Rosa H, Carvalho JP, Calado P, Martins B, Ribeiro R, Coheur L. Using fuzzy fingerprints for cyberbullying detection in social networks. In: *2018 IEEE International Conference on Fuzzy Systems (FUZZ)*. IEEE; 2018. p. 56-62.
- [17] Rosa H, de Matos DM, Ribeiro R, Coheur L, Carvalho JP. A “deeper” look at detecting cyberbullying in social networks. In: *2018 International joint conference on neural networks (IJCNN)*. IEEE; 2018. p. 1-8.
- [18] Heck RH, Thomas SL, Tabata LN. *Multilevel and longitudinal modeling with IBM SPSS*. Routledge; 2013.

- [19] Hox J. *Multilevel analysis: Techniques and applications* (2nd ed.). New York, NY: Routledge Academic; 2010.
- [20] Beheshti Z, Shamsuddin SMH. A review of population-based meta-heuristic algorithm. *Int J Adv Soft Comput Its Appl* 2013); 5:1–35.
- [21] Zhou Y, Wang F, Tang J, Nussinov R, Cheng F. Artificial intelligence in COVID-19 drug repurposing. *Lancet Digital Health* 2020;2:e667–76.
- [22] Fellous JM, Sapiro G, Rossi A, Mayberg H, Ferrante M. Explainable Artificial Intelligence for Neuroscience: Behavioral Neurostimulation. *Front in Neuroscience* 2019;13:1346.
- [23] Francisco SM, Veiga Simão AM, Ferreira PC, Martins MJD. Cyberbullying: The hidden side of college students. *Computers in Human Behavior* 2015;43:167–82.
- [24] Rachoene M, Oyedemi T. From self-expression to social aggression: Cyberbullying culture among south african youth on facebook. *Communication* 2015;41(3):302–19.
- [25] Norton RW. *Communicator style: Theory, applications, and measures*. Sage, Beverly Hills, CA; 1983.
- [26] Agarwal UA, Gupta M. Communication styles of millennials: Trends & relevance for the industry. *Indian Journal of Industrial Relations* 2018;53(3): 504–18.
- [27] Jakubowski P, Lange AJ. *The Assertive Option: Your Rights & Responsibilities*. Research Press Company; 1978.
- [28] del Pino CL, Sánchez A, Pérez-Nieto MA, Fernández M. Impulsividad, autoestima y control cognitivo en la agresividad del adolescente. *EduPsykhé. Revista de Psicología y Educación* 2008;7(1): 81–99.
- [29] den Hamer AH, Konijn EA. Can emotion regulation serve as a tool in combating cyberbullying? *Personality and Individual Differences* 2016;102:1–6.
- [30] Lin M. Risk factors associated with cyberbullying victimization and perpetration among Taiwanese children. PhD thesis, Texas Medical Center Dissertations (via ProQuest). AAI10126751; 2017.
- [31] Jacobs NCL, Dehue F, Völlink T, Lechner L. Determinants of adolescents' ineffective and improved coping with cyberbullying: A delphi study. *Journal of Adolescence* 2014;37(4):373–85.
- [32] Bridges LJ, Denham SA, Ganiban JM. Definitional issues in emotion regulation research. *Child Development* 2004;75(2):340–5.
- [33] Gross JJ. Emotion regulation: Affective, cognitive, and social consequences. *Psychophysiology* 2002;39(3):281–91.

- [34] Rottenberg J, Gross JJ. Emotion and emotion regulation: A map for psychotherapy researchers. *Clinical Psychology: Science and Practice* 2007;14(4):323–8.
- [35] Ochsner K, Bunge S, Gross J, and Gabrieli, J. Rethinking Feelings: An fMRI Study of the Cognitive Regulation of Emotion. *Journal of Cognitive Neuroscience* 2002;14(8):1215–29.
- [36] Mauss IB, Levenson RW, McCarter L, Wilhelm FH, Gross JJ. The tie that binds? coherence among emotion experience, behavior, and physiology. *Emotion* 2005;5(2):175–90.
- [37] Reisenzein R. A short history of psychological perspectives on emotion. In: *The Oxford Handbook of Affective Computing*, Oxford University Press, New York; 2015. pp. 1–34.
- [38] Scherer KR. Appraisal considered as a process of multilevel sequential checking. In: *Appraisal processes in emotion: Theory, Methods, Research*, Oxford University Press, Oxford; 2001. pp. 92–129.
- [39] Gross JJ, Thompson RA. Emotion regulation: conceptual foundations. In: *Handbook of emotion regulation*, Guilford Press; 2007.
- [40] Mauss IB, Cook CL, Gross JJ. Automatic emotion regulation during anger provocation. *Journal of Experimental Social Psychology* 2007;43 (5):698–711.
- [41] Mauss IB, Bunge SA, Gross JJ. Automatic emotion regulation. *Social and Personality Psychology Compass* 2007;1(1):146–67.
- [42] Taylor GJ, Bagby RM, Parker JDA. *Disorders of affect regulation: alexithymia in medical and psychiatric illness*. Cambridge University Press, Cambridge; 1997.
- [43] Gratz KL, Roemer L. Multidimensional assessment of emotion regulation and dysregulation: Development, factor structure, and initial validation of the difficulties in emotion regulation scale. *Journal of Psychopathology and Behavioral Assessment* 2004;26(1):41–54.
- [44] Gratz KL, Tull MT. Emotion regulation as a mechanism of change in acceptance- and mindfulness-based treatments. In: *Assessing mindfulness and acceptance: Illuminating the processes of change*, New Harbinger Publications, Oakland, CA; 2010.
- [45] Hofmann SG, Asmundson GJG. Acceptance and mindfulness-based therapy: New wave or old hat? *Clinical Psychology Review* 2008;28(1):1-16.
- [46] Robertson T, Daffern M, Bucks RS. Emotion regulation and aggression. *Aggression and Violent Behavior* 2012;17(1):72–82.

- [47] Van Cleemput K, Vandebosch H, Poels K, Bastiaensens S, DeSmet A, De Bourdeaudhuij I. The development of a serious game on cyberbullying. In: *Cyberbullying: from theory to intervention*. Routledge; 2015. pp. 93.
- [48] Paiva A, Dias J, Sobral D, Aylett R, Woods S, Hall L et al. Learning by feeling: Evoking empathy with synthetic characters. *Applied Artificial Intelligence* 2005;19(3-4):235–66.
- [49] Batson CD. These Things Called Empathy: Eight Related but Distinct Phenomena. In: *The Social Neuroscience of Empathy*, MIT Press, Cambridge, MA, USA; 2009. pp. 3–15.
- [50] Dinakar K, Jones B, Havasi C, Lieberman H, Picard R. Common sense reasoning for detection, prevention, and mitigation of cyberbullying. *ACM Transactions on Interactive Intelligent Systems* 2012;2(3):18:1–18:30.
- [51] Kowalski RM, Limber SP. Electronic bullying among middle school students. *Journal of Adolescent Health* 2007;41(6):S22–S30.
- [52] Olweus D. Bully/victim problems in school: Facts and intervention. *European Journal of Psychology of Education* 1997;12(4):495.
- [53] Olson EA, Collins PF, Hooper CJ, Muetzel R, Lim KO, Luciana M. White matter integrity predicts delay discounting behavior in 9- to 23-year-olds: a diffusion tensor imaging study. *J Cogn Neurosci* 2009;21:1406–21.
- [54] Crone EA, Bullens L, van der Plas EA, Kijkuit EJ, Zelazo PD. Developmental changes and individual differences in risk and perspective taking in adolescence. *Dev Psychopathol* 2008;20:1213–29.
- [55] Bava S, Boucquey V, Goldenberg D, Thayer RE, Ward M, Jacobus J, et al. Sex differences in adolescent white matter architecture. *Brain Res* 2011;1375:41–8.
- [56] Linacre JM. Winsteps® Rasch measurement computer program User's Guide. Winsteps.com, Beaverton, Oregon. Retrieved from <https://www.winsteps.com/manuals.htm>.
- [57] Ferrando P. Assessing short-term individual consistency using IRT-based statistics, *Psicológica* 2010;31:319-34.
- [58] Rasch G. Probabilistic models for some intelligence and attainment tests. Chicago. IL: University of Chicago Press; 1980. (original work published 1960)
- [59] Wright BD, Masters GN. Rating scale analysis. Chicago, IL: Mesa Press; 1982.
- [60] Bond TG, Fox CM. Applying the Rasch model: Fundamental measurement in the human sciences. Erlbaum, Mahwah, NJ; 2007.

- [61] Smith Jr EV. Evidence for the reliability of measures and validity of measure interpretation: a rasch measurement perspective. *Journal of Applied Measurement* 2001;2(3):281–311.
- [62] Fox CM, Jones JA. Uses of rasch modeling in counseling psychology research. *Journal of Counseling Psychology* 1998;45(1):30–45.
- [63] Rosa H, Pereira N, Ribeiro R, Ferreira PC, Carvalho JP, Oliveira S et al. Automatic cyberbullying detection: A systematic review. *Computers in Human Behavior* 2019;93:333–45.
- [64] Zimmerman BJ. From cognitive modeling to self-regulation: A social cognitive career path. *Educational Psychologist* 2013;48(3):135–47.
- [65] Bandura A. *Aggression: a social learning analysis*. Prentice-Hall, Englewood Cliffs, NJ; 1973.
- [66] Kirchner EP, Kennedy RE, Draguns JG. Assertion and aggression in adult offenders. *Behavior Therapy* 1979;10(4):452–71.
- [67] Hinduja S, Patchin JW. *Bullying beyond the schoolyard: Preventing and responding to cyberbullying*. Sage Publications, Thousand Oaks, CA; 2009.
- [68] Hinduja S, Patchin JW. Bullying, cyberbullying, and suicide. *Archives of Suicide Research* 2010;14(3):206–21.
- [69] Bono R, Blanca MJ, Arnau J, Gómez-Benito J. Non-normal Distributions Commonly Used in Health, Education, and Social Sciences: A Systematic Review. *Frontiers in Psychology* 2017;8:1602.
- [70] Marôco J. *Análise Estatística com o SPSS Statistics (7th ed.)*. Pêro Pinheiro: ReportNumber; 2018. Retrieved from [www.reportnumber.pt/ae](http://www.reportnumber.pt/ae)
- [71] McCoach DB. Hierarchical linear modeling. In: *The reviewer's guide to quantitative methods in the social sciences*, Routledge, New York, USA; 2010. pp. 123–40.
- [72] Singer JD, Willet JB. *Applied longitudinal data analysis: Modeling change and event occurrence*. Oxford University Press, New York, USA; 2003.
- [73] Weissberg R, Durlak J, Domitrovich C, Gullotta TP. Social and emotional learning: Past, present, and future. In: *Handbook of social and emotional learning: Research and practice*, Guilford Press, New York, NY, US; 2015. pp. 3–19.
- [74] Ferreira PC, Veiga Simão AM, da Silva AL. Does training in how to regulate one's learning affect how students report self-regulated learning in diary tasks? *Metacognition and Learning* 2015;10(2):199–230.

## 11. Tables

Table 1

Description of resources to foster seeking information strategies.

Resource	Description
1. Did you know?	Informational content which refers to trivia about bullying and cyberbullying, and the use of social networks (e.g., Did you know that, whether you were the author or not of a given publication, if you share it in your profile, it makes you responsible for it?).
2. Know more about	Informative content on cyber security strategies, such as passwords and sharing information online (e.g., Use online privacy policies. Create restrictions for whom has access to your information.), and cyberbullying prevention and intervention (e.g., Log out in your profiles! Regardless of how convenient it is to save logins, you never know who can access them besides you).
3. True stories	Narratives of cyberbullying events adapted from testimonies of adolescents which include strategies to positively deal with the situation (e.g., blocking the aggressor, reporting the situation).
4. Videos	Informative content about cyberbullying and testimonials (e.g., <a href="https://www.youtube.com/watch?v=f6K9Ie_Chjs">https://www.youtube.com/watch?v=f6K9Ie_Chjs</a> ).

Table 2

Descriptive statistics and fixed and random effects parameter estimates for model predicting application use.

Application Use	Level 1 ( <i>N</i> = 156)			Level 2 ( <i>N</i> = 39)		
Sessions	0	1	2	3		
	M(SD)	M(SD)	M(SD)	M(SD)	M(SD)	
Gender	Female	46.72(84.31)	53.63(46.31)	37(39.19)	36(36.52)	43.34(53.60)
	Male	28.10(59.22)	48.71(68.22)	52(47.35)	62.82(36.59)	47.91(54.91)
Parameter	Intercept-only	Intercept +Time	With Predictors			
Fixed Effects						
Intercept	46.62**(5.64)	30.78**(10.11)	55.54**(18.78)			
Time		6.33 <sup>†</sup> (3.35)	-4.88 (6.20)			
Gender			-34.49(22.17)			
Time*Gender			15.62*(7.31)			
Random Effects						
Repeated measures						
Intercept	2263.70**(295.96)	2196.80**(287.21)	2114.40**(276.44)			
Intercept	678.72*(291.40)	695.44*(290.85)	711.81*(289.27)			
Deviance	1678.50	1674.99	1670.39			
AIC	1684.50	1682.99	1682.39			
BIC	1693.65	1695.19	1700.69			

*Note:* The Level 1 means and standard deviations are reported according to the time variable (from 0 to 3).

Standard errors are in brackets. † $p < 0.10$ ; \* $p < 0.05$ ; \*\* $p < 0.01$ .

Table 3

Percentage of time adolescents dedicated to specific resources to guide them towards pro-social behavior.

Resources	Female	Male	
Introductory Communication Practice Exercise	7.8%	8%	
Self-reactiveness	Express Yourself	11.3%	11.9%
	Inspire Yourself	9.7%	17.6%
Seeking Information	Did you know?	27%	24.7%
	Know more about	12.2%	8.1%
	True stories	30.2%	28.1%
	Videos	1.5%	1.4%
Seeking Social Assistance	Links	0.2%	0.3%

## 12. Figures

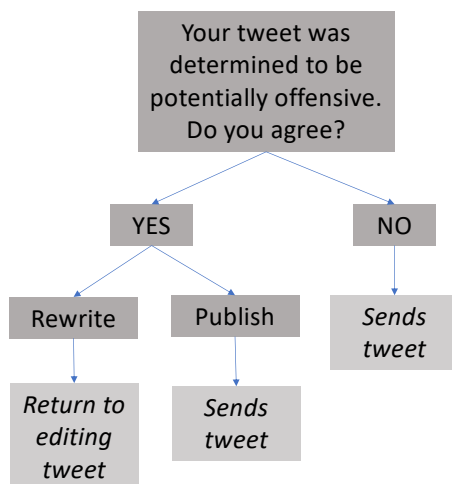


Figure 1: ComViver Online’s chain of options triggered by an alert to senders.

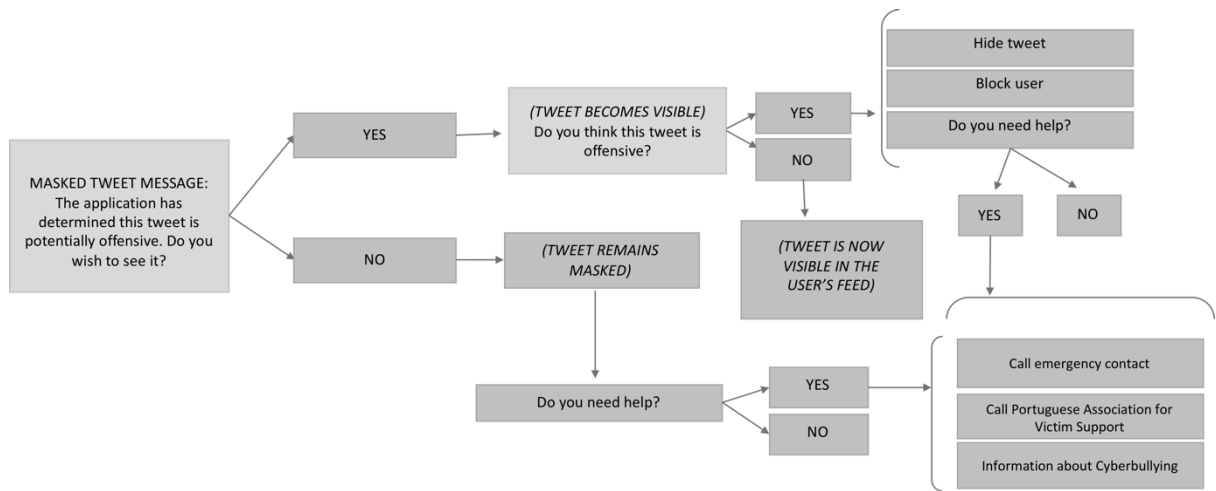


Figure 2: ComViver Online's chain of options triggered by an alert to receivers.



Figure 3: Example from the resource Inspire Yourself.

Which of these responses do you think would be the best to express this thought?

- What did I do wrong?
- I'm going to punch you in the face!
- I won't accept you threatening me.

Figure 4: Example from the resource Express Yourself.

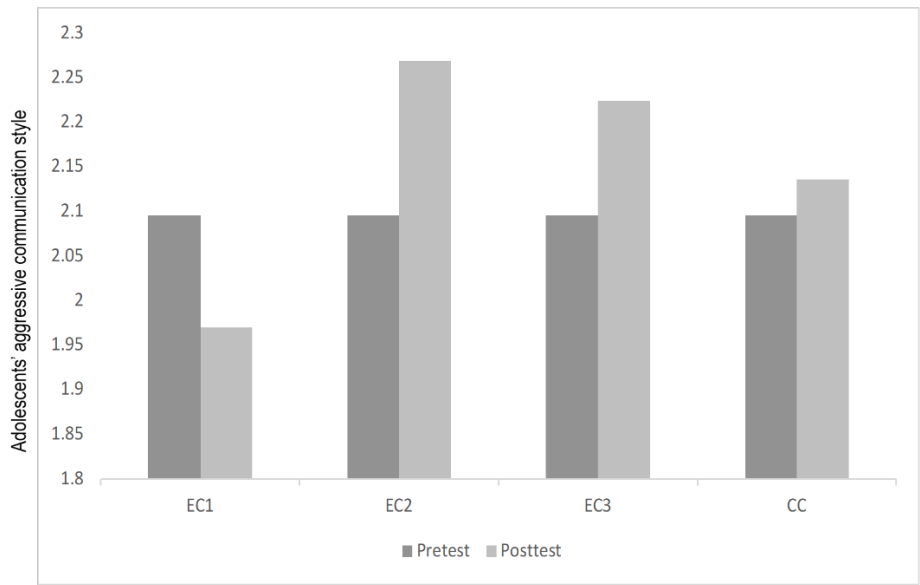


Figure 5: Estimated marginal means of adolescents' aggressive communication style in the pretest and posttest.

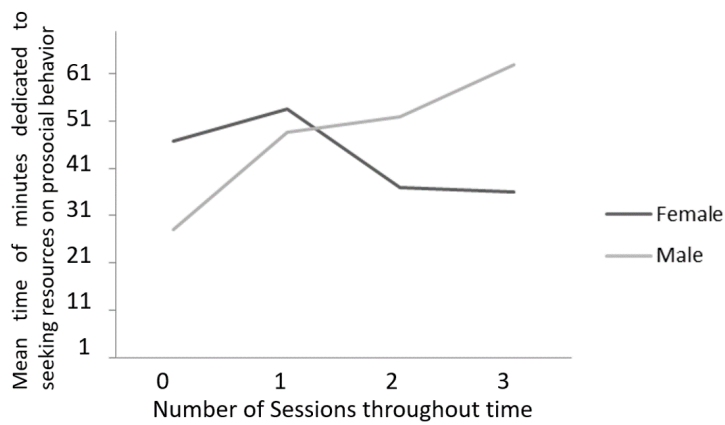


Figure 6: Male and female participants' trajectories throughout the different sessions with regards to the time they dedicated to seeking resources on prosocial behavior in the application.