# Self-Disclosure of Bullying Experiences and Social Support in Avatar Communication: Analysis of Verbal and Nonverbal Communications

**Masanori Takano,**[1] **Takaaki Tsunoda**[1]

[1]CyberAgent, Inc.

takano_masanori@cyberagent.co.jp

## Abstract

Avatar communication through the Internet has great potential to be an appropriate environment for self-disclosure and social support. Anonymity and ease of access drive self-disclosure of even the most serious problems. Rich nonverbal communication, co-presence, and real-time interaction increase emotional closeness. However, there has not been much research with regard to examining social support in avatar communication. In this paper, we aim to facilitate self-disclosure and social support for bullied people through avatar communication. For this purpose, we analyzed verbal and nonverbal communication about bullying experiences through an avatar communication service. We demonstrate that people who emotionally disclosed their bullying experiences received better social support. In addition, people who provided social support used emotional expressions to convey emotional empathy. These were observed in conversations with a few acquaintances in closed spaces. Our findings reveal areas where we can improve upon the design of avatar communication spaces for effective social support.

People support others, and they are also supported by others. This social support, such as showing empathy, affection, and respect, improves people's mental health (Turner, Turner, and Hale 2014). This behavior also works through the Internet (online social support). People who provide online social support are not necessarily closely connected to the recipients of the support in the physical world. Online social support therefore has great potential to act as an auxiliary for social relationships in the physical world (Chung 2013). For example, social networking sites' (SNS) usage of adolescent mothers enables access to tangible, informational, and emotional support (Nolan, Hendricks, and Towell 2015). On the bulletin board website Reddit, users provide attentive responses and empathize with users who disclose their experiences of sexual abuse (Andalibi et al. 2016), suicidal feelings (De Choudhury and De 2014), and mental health problems (Sharma and De Choudhury 2018). In a massively multiplayer online game (MMO), game players feel that other players will help them, and this feeling increases their well-being (Cole and Griffiths 2007; Kaczmarek and Drążkowski

Figure 1: Communications in Pigg Party. Users visit rooms of friends and strangers to chat with them through their avatars.

2014). On the Q&A site ASKfm (Ashktorab et al. 2017) and the photo-sharing site Flickr (Wang et al. 2017), some users require social support of many and unspecified users.

People have to resort to self-disclosure to receive social support from others. Disclosing themselves in detail facilitates the reception of greater social support in cases of negative experiences, negative emotions, and suicidal feelings (De Choudhury and De 2014). However, disclosing serious problems, such as bullying experiences and sexual abuse experiences, is often difficult because of the fear of rejection by listeners (Omarzu 2000; Andalibi et al. 2016).

The anonymity of the Internet drives self-disclosure about negative experiences (Kang, Dabbish, and Sutton 2016). For example, sexual minorities who tended to be unacceptable by society at that time of publication of the paper (McKenna and Bargh 1998) recovered self-esteem through anonymous communication communities on the Internet (newsgroups) (McKenna and Bargh 1998). Andalibi et al. (2016) showed that anonymity drives self-disclosure of sexual abuse on Reddit. On ASKfm, where users can post anonymously, users did self-disclosure of experiences of self-harm and suicidal feelings (Ashktorab et al. 2017). Interestingly, some of them posted questions to themselves anonymously about their experiences with self-harm and

suicidal feelings for self-disclosures. On Flickr, users have been found to self-harm photos in expectation of receiving care, empathy, and solicitude from others (Wang et al. 2017). Additionally, anonymous self-disclosure facilitates self-disclosure in the physical world (McKenna and Bargh 1998; Liu and Yu 2013; Andalibi and Forte 2018).

Nonverbal communication is also important for self-disclosure and social support (Mehrabian 1970; Manusov and Patterson 2006). Especially, avatar communication, where people with virtual bodies can show facial expressions and gestures in virtual space, allows us to nonverbally interact through the Internet (Antonijevic 2008; Green-Hamann, Campbell Eichhorn, and Sherblom 2011) while maintaining anonymity (Küster, Krumhuber, and Kappas 2015). For example, proxemic behavior in the online virtual world game Second Life has been observed to provide feelings similar to those in the physical world (Antonijevic 2008). In an MMO, nonverbal communication drives cooperation between players (O'Connor et al. 2015). These features of avatar communication provide better social support (Green-Hamann, Campbell Eichhorn, and Sherblom 2011). Green-Hamann et al. (2011) showed that members of social support groups in Second Life (Alcoholics Anonymous and Cancer Caregivers) prefer nonverbal communication caused by co-presence, real-time interaction, and proxemic behavior. This nonverbal communication is important for emotional communication including self-disclosure and social support (Mehrabian 1970; Manusov and Patterson 2006). In the groups where the members constructed close relationships, they did self-disclosure and provided emotional social support to each other, i.e., alcoholics and their family's serious diseases (Green-Hamann, Campbell Eichhorn, and Sherblom 2011).

Although these features of avatar communication drive social support between users (Green-Hamann, Campbell Eichhorn, and Sherblom 2011), examining social support in an avatar communication space has barely received any attention. In this paper, we focus on self-disclosure of people who were bullied in the physical world (almost all were students in schools) and social support for them in a Japanese avatar communication service (Figure 1). In contrast to approaches of previous studies (research interviews (Green-Hamann, Campbell Eichhorn, and Sherblom 2011; O'Connor et al. 2015), questionnaire surveys (Cole and Griffiths 2007), and ethnographic researches (Antonijevic 2008)), we quantitatively analyze actual utterances and avatar actions in avatar communication using user conversation and action logs. This analysis of detailed behavior data enables an approach to studying the time evolution of bullying experience conversations for designing avatar communication spaces, which support users in real-time, e.g., avatar action suggestion and providing triggers for self-disclosure.

The reason for using bullying experience conversation data is that we expect that online social supports, especially in avatar communication services providing physical-world-like communication, supplement the existing support for bullied children. Communication tools over the Internet, which are popular with young people, decrease bullied children's hesitancy in accessing support. For example, bullied

Japanese children ask for help from school counselors using the famous messenger application LINE[1] significantly more often than by calling telephone call centers (The Japan Times 2017).

Our research questions are as follows:

- **RQ1:** Can self-disclosure in avatar communication initiate social support for bullied people?

- **RQ2:** What should bullied people disclose for getting effective social support? How do listeners reply to these people to ensure effective social support?

- **RQ3:** How do avatar actions (gestures) support their self-disclosure and social support?

The summary of this paper reveals the following findings:

- **RQ1:** Bullied people get effective social support when they do self-disclosure about their bullying experience to a few acquaintances in closed spaces.

- **RQ2:** Bullied people, who give an explanation of bullying experiences and their emotional damage, receive effective social support. Listeners discuss bullying and provide empathize with them.

- **RQ3:** Avatar actions appear to facilitate sensitive bullying experience conversations and their effects. These support their verbal emotional disclosure and empathy.

In addition, we show the features of communication in unestablished self-disclosure and social support. Our findings contribute to the improvement of avatar communication tools through the Internet for effective social support for bullying experiences. We discuss improving avatar communication services for self-disclosure and social support for bullied people based on the findings.

## Materials and Methods

### Avatar Communication Service

In this study, we use conversation data from *Pigg Party*[2], a Japanese avatar communication service shown in Figure 1. The data set comprises utterance text, avatar actions, and types of communication spaces (room types that we describe later). These were extracted from user conversations and action logs provided by the Pigg Party application provider.

Pigg Party users can communicate synchronously with their own designed avatar in virtual spaces. In addition to sending a text message, users can respond with dozens of avatar animations we call *avatar actions*, as shown in Figure 2. Regarding demographics, 61% of the users are teenagers, and 65% of users are females.

Although conversation topics are not limited to certain subjects, there are several conversations related to school life, including conversations related to bullying, as more than half of the users are teenagers. Within the service, users make social relationships that are different from their real life (Takano and Mizuno 2018), that is, users tend not to talk with their families or physical world friends. This fact is the

---

[1]https://line.me/en/

[2]https://lp.pigg-party.com. Available for iOS and Android devices.

notable difference from major SNS in Japan, such as Facebook, Twitter, and LINE (MIC 2017).

Pigg Party offers three types of communication spaces we call *rooms*: **private**, **public**, and **temporary** rooms.

- **Private room**: Each user owns this type of room. Users can enter a private room by the following ways: **1)** by clicking the *enter* button shown in a room owner's profile window[3], or **2)** by random entry, which is initiated by a random entry mode. The capacity of a room is 10 users.

- **Public room**: This room is offered as a public space that anyone can enter. The capacity of a room varies between rooms. The largest room has up to 20 users.

- **Temporary room**: This room can be an ad hoc space for any user. Anyone can enter an arbitrary room found in a temporary room search page. The capacity of this room type is 10 users. This room type is closed two hours after it is created.

Note that the Pigg Party users accept the terms of use and privacy policy, which allow the analysis of their conversations and behavior data for service improvements. In addition, when a user starts conversations for the first time, this service notifies that the Pigg Party application provider monitors user conversations. The data set is pseudonymized. Furthermore, we use the data set after removing identifying words for de-identification.

## Bullying Experience Conversation Data

We define the terms in this paper as follows. **Bullying experience conversations** as conversations involving users talking about their bullying experiences. **Bullied users** as users bullied at that time or refusing to go to school due to bullying[4]. **Listeners** as participants in bullying experience conversations, excluding bullied users. We regard a bullied user and a listener as **acquaintances** if they have communicated within 30 days prior to the bullying experience conversation. Otherwise, they are regarded as **non-acquaintances**.

We identified 1,130 bullying experience conversations from conversations in the period from September 2016 to August 2017 through manual annotation as follows. First, we extracted conversations including the word "bullying" or its homonyms (Fig. 3) from all conversations in the period from September 2016 to August 2017. Next, we picked utterances within a 30-min window that is from 15-min before an utterance including "bullying" to 15-min after the utterance (if a conversation had multiple "bullying" words, a window was from 15-min before the first one to 15-min after the last one). Avatar actions in the window were extracted as well. Afterward, we randomly sampled 40,000 conversations and manually annotated whether a conversation was a bullying experience conversation. Accordingly, we identified 1,130 bullying experience conversations. These conversations have 287,726 utterances, 1,025 bullied users, and 6,613 listeners. Regarding the room type, 527 private rooms, 41 public rooms, and 562 temporary rooms were used for a

---

[3]Another user's profile window can be viewed if the user is a *acquaintance* or is in the same room.

[4]We did not include cyberbullying as scope of our work.



Figure 2: Examples of avatar actions, named *laugh* (left), *sad* (center), and *angry* (right). Pigg Party offers dozens of actions besides these examples.

"いじめ" "イジメ" "苛め" "虐め"

Figure 3: Homonyms of "bullying" in Japanese.

bullying experience conversation. For text preprocessing, we applied the Japanese morphological analyzer MeCab (Kudo et al. 2004)[5] and selected content words, i.e., nouns, verbs, adverbs, and emojis.

Private rooms had a few users (median is 3) compared to private and temporary rooms (both medians are 11). Although the capacity of a temporary room is 10 users, some rooms had more than 10 users. The reason for this is that many users enter and leave one after another rapidly in these rooms.

A significant amount of listeners (54.4%) are acquaintances with bullied users in private rooms, while most of the listeners are non-acquaintances in public and temporary rooms (acquaintance ratios are 10.0% and 23.5%, respectively) [6]. This is likely attributed to a difference in the ways to enter rooms.

## Effects of Social Support

In this section, we describe how to evaluate the effects of social support on bullied people and listeners using user behavior logs. For this purpose, we evaluate usage frequencies of users of Pigg Party to measure changes in user satisfaction before or after having bullying experience conversations as the effects of social support. These usage frequencies correlated user satisfaction (Rokito et al. 2019). That is, we consider that if a bullying experience conversation gives a positive impression to users, then they will use Pigg Party more frequently.

We evaluated the effect of bullying experience conversations on user usage frequencies by the following generalized linear model (GLM):

$$y \sim B(7, p) \tag{1}$$
$$\text{logit}(p) = \beta_1 f + \beta_2 b_r + \beta_3 b_u + \beta_4 b_t$$

---

[5]We employed the mecab-ipadic-NEologd dictionary (https://github.com/neologd/mecab-ipadic-neologd) as the MeCab dictionary.

[6]The comparison of acquaintance ratios between private rooms and others showed a significant difference at $p$-value $< 2.2 \times 10^{-16}$, chi-squared test.

Table 1: Explanatory variables of GLM (Eq. 1)

| Variable | Description |
|---|---|
| $f$ | Login days in a week having the bullying experience conversation, $\{0, \ldots, 7\}$ |
| $b_r$ | Whether the user talked about their bullying experiences in a private room, $\{0, 1\}$ |
| $b_u$ | — in a public room, $\{0, 1\}$ |
| $b_t$ | — in a temporary room, $\{0, 1\}$ |
| $c_r$ | Whether the user listened to a bullying experience in a private room, $\{0, 1\}$ |
| $c_u$ | — in a public room, $\{0, 1\}$ |
| $c_t$ | — in a temporary room, $\{0, 1\}$ |

$$+\beta_4 c_r + \beta_5 c_u + \beta_6 c_t +$$
$$+ \sum_{w_i} \gamma_w w_i + \beta_0$$

where $y$ is the number of login days in a week after having a bullying experience conversation, $B$ is the binomial distribution, $p$ is the probability of the user's logins in a day, and $w_i$ denotes the time-varying covariates for each week $i$ ($1 \leq i \leq 52$). The other explanatory variables $f$, $b$, and $c$ are described in Table 1.

## Analysis of Conversation Texts

We extracted characteristic topics of bullying from conversations, including bullying experiences, bullying environments, and insults by topic modeling (latent Dirichlet allocation; LDA (Blei, Ng, and Jordan 2003)). Previous works (De Choudhury and De 2014; Ashktorab et al. 2017; Jaidka, Guntuku, and Ungar 2018; Ernala et al. 2018) used a topic model (LDA) for understanding texts related to self-disclosure and/or social support to be an efficient way of discovering topics automatically, organize, and categorize large amounts of text.

Prior to applying a topic model, we incorporated non-bullying experience conversations into a bullying experience conversation data set. The reason for this is that our data set had only bullying experience conversations and applying a topic model to this fails to extract topics related to bullying. Specifically, in addition to 1,130 bullying experience conversations, 3,000 conversations randomly selected from the same period as bullying experience data set were used.

To extract latent topics, the LDA was employed. We regarded conversations in a 30-min window as a document for LDA. We set LDA parameters as the number of topics $K = 200$ and $\alpha = 50/200$, respectively.

Characteristic topics were selected in bullying experience conversations after applying the LDA, on the basis of the pointwise mutual information (PMI) score. PMI is defined as follows:

$$\text{PMI}_b(t) = \log_2 P(t|b) - \log_2 P(t) \tag{2}$$

where $t$ is a topic and $b \in \{0, 1\}$ is the bullying experience conversation flag. PMI allows us to select topics distinguishing between bullying experience conversations and others,
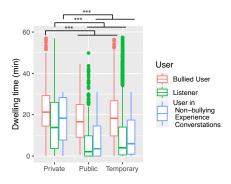


Figure 4: Dwelling time of bullied users, listeners, and users in non-bullying experience converstations by room types. *** indicates a significant difference at $p$-value $< 0.001$, Wilcoxon's rank-sum test.

since $\text{PMI}_b(t) > 0$ denotes the topic $t$ characteristic under the given condition $b$. On the basis of PMI, we first selected topics that had $\text{PMI}_{b=1}(t) > 0$.

Finally, we removed topics that were not related to bullying, bullied users, bullying environments, or insults. Here, we included insult topics to see if disclosing bullying experiences elicited insults from listeners (like in ASKfm (Ashktorab et al. 2017)).

## Results
### Effects of Social Support (RQ1)

We first evaluate the effectiveness of social support by room types used for bullying experience conversations.

Table 2 reports the results of the regression analysis. The explanatory variables $f, b_r, b_u, c_r, c_t,$ and some $w_i$ were chosen by Akaike's information criterion (AIC). For the effects on bullied users, talking in private rooms had positive effects on the usage frequency ($b_r > 0$). In contrast, talking in public rooms worked negatively ($b_u < 0$). Talking in temporary rooms does not affect usage frequencies much, i.e., AIC did not choose $b_t$. Listening in private or temporary rooms had positive effects ($c_r, c_t > 0$).

The dwelling time of users also indicates the difference among the room types. Figure 4 shows users' dwelling time by room types. As shown in the figure, users in private rooms stayed for a longer time compared to other rooms for all types of users including bullied users.[7] This result suggests that they conversing in private rooms had better experiences compared to users talking in other rooms.

The results imply that bullied users received better social support in private rooms. Conversely, self-disclosure and/or

---

[7] Note that it is hard to compare the dwelling time in bullying experience conversations with that in non-bullying experience conversations because the data extraction settings of the latter which did not have the data gathering keywords ("bullying" and its homonyms) were different with the former. We extracted utterances within a 30-min window from randomly selected utterances for constructing the data-set of non-bullying experience conversations.

Table 2: Coefficients of GLM (Eq. 1). Note that $b_t$ and $c_u$ are not chosen by AIC. Covariates $w_i$ are not listed here as they are not essential to the analysis. Var., Coef., and S.E. are variable, coefficient, and standard error, respectively. ***, **, and * indicate that the signs of regression coefficients did not change in a Wald-type $99.9\%$, $99\%$, and $95\%$ confidence interval.

| Var. | Coef. | S.E. | $t$-value | $p$-value |
|---|---|---|---|---|
| $f$ | 0.504 | 0.003 | 180.664 | $< 2.0 \times 10^{-16}$ *** |
| $b_r$ | 0.287 | 0.041 | 7.003 | $2.5 \times 10^{-12}$ *** |
| $b_u$ | $-0.282$ | 0.131 | $-2.145$ | $3.2 \times 10^{-2}$ * |
| $c_r$ | 0.267 | 0.026 | 10.068 | $< 2.0 \times 10^{-16}$ *** |
| $c_t$ | 0.103 | 0.014 | 7.470 | $8.0 \times 10^{-14}$ *** |
| Intercept | $-1.109$ | 0.043 | $-25.838$ | $< 2.0 \times 10^{-16}$ *** |

social support are observed to be lacking in temporary and public rooms. This analysis suggests that there is effective social support for bullied people with a few acquaintances in closed spaces (private rooms). In the following sections, we focus on chat texts and avatar actions to explore why only conversations with a few acquaintances in closed spaces were effective for bullied users.

## Verbal Communication (RQ2)

We extracted characteristic topics of bullying from conversations, including bullying experiences, bullying environments, and insults, using a topic model. Accordingly, 11 topics characterizing bullying experience conversations were extracted as shown in table 3. These are six topics for bullying and pain (Bullying, Self-Body Shaming, Emotional Damage, Family Strife, School Refusal, and Suicide), three topics for schools (Class, School, and Sports), and two topics for insult exchange (Aspersion and Provocation). Table 3 lists the topics with their typical phrases.

Figure 5 shows the topic frequency of each room and each user type in three time periods (from $-15$ min to 0 min, from 0 min to 15 min, and from 15 min to 30 min, where the origin (0 min) is the first time of the utterance of the word "bullying" or its homonyms).

**Conversation Topics of Self-Disclosure** In all room types, bullied users self-disclosed of their bullying experiences. That is, they explained getting bullied (the Bullying topic; e.g., suffering violence and considering going for counseling; Figure 5a) and they talked about their classmates (the Class topic; e.g., hope for a new class in the next term; Figure 5b). In addition, they disclosed very sensitive topics (Family Strife, Self-Body Shaming, School Refusal, and Suicide topics; Figure 5a).

Especially in private rooms, bullied users self-disclosed more in-depth than in other room types. They talked about Emotional Damage more frequently than in other rooms.

**Conversation Topics of Response by Listeners** In private and temporary rooms, listeners responded according to self-disclosures of bullied users. That is, conversation topics of listeners are correlated with topics of bullied users (Bullying and Class in both room types and Emotional Damage in private rooms; Figure 5a, b). The high frequency of the

Emotional Damage topic suggests that listeners showed empathy to bullied users in private rooms (Figure 5a).

On the contrary, in public rooms, almost all listeners pay no attention to the self-disclosure of bullied users. Listeners in public rooms did not talk about bullying-related topics (Figure 5a). This is apparently the reason for the negative effects of self-disclosure in public rooms (Table 2).

In temporary rooms, bullied users and listeners sometimes insulted each other. The frequencies of the Aspersion and Provocation topics were high and increased with time in these rooms (Figure 5c).

## Avatar Actions (RQ3)

We analyzed avatar actions (facial expressions and gestures) in bullying experience conversations. Public rooms are excluded in the following analysis because users did not discuss bullying experiences in these rooms.

Figure 6 shows frequencies of actions per person. Overall, bullied users and listeners in private rooms used avatar actions more frequently than those in temporary rooms.

Bullied users tended to also use negative expressions (Sad, Terrified, Wailing) in both room types. This shows that they expressed painful emotions. Interestingly, bullied users in the last period (15–30 min) of private rooms tended not to use negative expressions. Instead, they tended to use positive expressions (Loud laugh, Laugh, Bow, Funny, and Happy). This suggests that bullied users in private rooms had better experiences compared to bullied users in temporary rooms.

Bullied users and listeners tended to use smiling expressions frequently, i.e., Laugh and Loud Laugh actions (like lol), especially in private rooms. Jokes were cracked near these actions in bullying experience conversations.

## Discussion

### Established Self-Disclosure and Social Support

We found that bullied users who received better social support did self-disclosure of bullying experiences emotionally, including verbal and nonverbal expressions. In addition, users who provided social support used emotional expressions to convey positive feelings and empathy. These were facilitated in conversations with a few acquaintances in closed spaces (private rooms; *RQ1*).

Table 3: Characteristic topics of bullying with their typical phrases

| Topic | Phrases |
|---|---|
| Bullying | bullying, bully, prejudice, disliked person, violence, kick, unfair fight, hyperpnea, unclothe, depression, piss pot, counselor, children's welfare center, board of education |
| Emotional Damage | taunt, deceived, kick, distress, grueling, cutting-off, trust, restoration of friendship |
| Family Strife | domestic violence, blap, worst feelings, children's welfare center, brother, individual enrichment courses |
| Self-Body Shaming | weight, height, fat, pipsqueak, scrawny, complex, dieting, brassiere |
| School Refusal | school refusal, social withdrawal, low education, repeat a school year, misanthropic, alternative school, NEET |
| Suicide | suicidal feelings, death-leap, hurt, hanging, (my parents will) feel sadness, survive, life, kill, depression, wrist, cutting knife, mommy |
| Class | next term, new class, changing schools, homeroom teacher, junior high school, withdrawal, disconnectedness, wallflower, escalate |
| School | school, examination score, club, casual talk, train |
| Sports | swimming, talented in sports, baseball, childish |
| Aspersion | disabled person, chimp, intelligence, bad, hypocritical person, low education, lol |
| Provocation | criticism, insult exchange, confute, ego trip, unamused, rights, common sense, intelligence, problematic kid, benefits |

Bullied users who received effective social support expressed not only information about their bullying experiences but also their emotions, such as painful feelings, to listeners. The content of their self-disclosure is as follows: i) explanations of bullying experiences (the Bullying topic; *RQ2*), ii) disclosure of very sensitive topics (the topics of Family Strife, Self-Body Shaming, School Refusal, and Suicide; *RQ2*), iii) verbal and rich nonverbal expressions for painful emotion (the Emotional Damage topic and the Wailing action; *RQ3*), and iv) frequent usage of laughing nonverbal expressions such Loud Laugh and Laugh, often with jokes (not only bullied users but also listeners; *RQ3*).

These show that findings in previous studies of online self-disclosure, which showed that self-disclosure tends to include emotional expressions and suicidal feelings (Reddit (De Choudhury and De 2014; Andalibi et al. 2016), Instagram (Andalibi, Ozturk, and Forte 2017), and Flickr (Wang et al. 2017)), are also realized in avatar communication.

In addition, in avatar communication, we observed that self-disclosures were probably driven by a communication interface showing facial expressions and gestures through avatars, and exclusiveness of private rooms. Nonverbal communication such as facial expressions and gestures increases closeness and satisfaction of communication (LaCrosse 1975; Koyama et al. 2017; Vlahovic, Roberts, and Dunbar 2012). This tends to be limited in text communication (most of non-avatar online communication) (Antonijevic 2008; Green-Hamann, Campbell Eichhorn, and Sherblom 2011), e.g. emojis and emoticons (Kralj Novak et al. 2015; Aldunate and González-Ibáñez 2017). People tend to disclose their emotion and sensitive information in closed cyber spaces with friends (not strangers) (Jaidka, Guntuku, and Ungar 2018). Consequently, bullied users may have disclosed their bullying experiences and very sensitive topics with their emotions. These disclosures of their emotions and sensitive topics facilitate social support (De Choudhury and De 2014).

Further, in private rooms, bullied users and listeners tended to tell jokes with laughing actions to each other. Telling jokes to each other conveys positive feelings and facilitates self-disclosure (Ervin-Tripp and Lampert 2009).

Listeners who provided social support did the following things in response to the self-disclosure of bullied users. **a)** Discussion of bullying (the Bullying topic; *RQ2*), and **b)** empathy for bullied users' painful emotion by verbal and nonverbal expressions (the Emotional Damage topic and the Wailing action; *RQ2 and 3*). This empathy for painful emotions, known as emotional social support, improves receivers' mental health (Turner, Turner, and Hale 2014). The emotional support has also been reported by previous works for online social support (Reddit (Andalibi et al. 2016; De Choudhury and De 2014), Facebook (Nolan, Hendricks, and Towell 2015; Andalibi and Forte 2018), and Instagram (Andalibi, Ozturk, and Forte 2017)).

Social support also positively affects listeners who provided social support as with the physical world (Morelli et al. 2015; Inagaki et al. 2016). Interestingly, it was observed not only when social support effectively worked (private rooms) but also when social support did not work (temporary rooms).

## Unestablished Self-Disclosure and Social Support

In spaces where listeners were mostly strangers (non-acquaintances in public and temporary rooms), the positive effect of social support was not observed despite bullied users disclosing their bullying experiences.

In both room types, bullied users tend not to mention their own painful emotions (the Emotional Damage topic) in spite of disclosure of their bullying experiences. The reason for this apparently is that people tend to disclose in-depth in

(a) Topics for pain



(b) Topics for schools
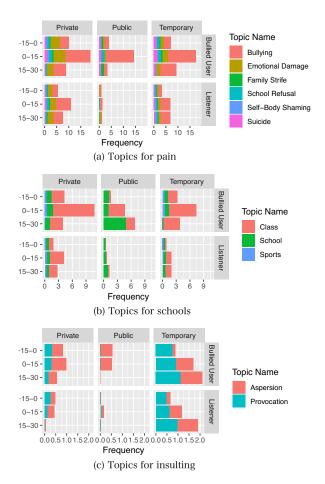


(c) Topics for insulting

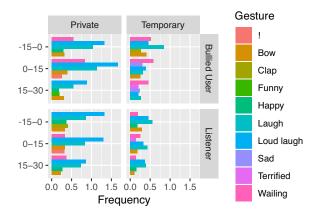Figure 5: Frequencies of bullying-related topics



Figure 6: Frequencies of avatar actions

spaces limiting relationships (Jaidka, Guntuku, and Ungar 2018).

In contrast to our result, previous works showed that social support also worked effectively in open spaces with asynchronous communication, such as Reddit (De Choudhury and De 2014; Andalibi et al. 2016; Sharma and De Choudhury 2018) and Instagram (Andalibi, Ozturk, and Forte 2017). This difference is apparently caused by the differences between synchronous and asynchronous communications as follows.

The first reason is that social support given by listeners did not work effectively because bullied users and listeners insulted each other in some conversations in temporary rooms. These conversations had many passing strangers who sometimes engaged users in a hurtful manner. Such cyberbullying is a serious problem on the Internet, e.g., in Reddit (Andalibi et al. 2016) and ASKfm (Ashktorab et al. 2017). This problem in synchronous communication is more fatal than in asynchronous communication because it may be difficult not to mind the insults in synchronous communication.

The second reason is that listeners did not pay attention to bullied users' self-disclosure in public rooms. It is apparently because utterances in synchronous communication caught less attention of listeners because the utterances became invisible in seconds. This negatively affected bullied users in this room type, which also included many strangers.

## Facilitating Self-Disclosure and Social Support

In this section, we discuss improving avatar communication services for self-disclosure and social support for bullied people based on our analysis results.

First, facilitating self-disclosure when in conversations with a few friends in closed spaces by avatar communication systems based on the results of RQ1, e.g., providing conversation topics as triggers for self-disclosure in these conversations. Providing the triggers appears to improve user experiences of not only users who disclose themselves but also users who provide social support because providing social support increases providers' well-being (Morelli et al. 2015; Inagaki et al. 2016). In ASKfm, some of the users asked themselves under anonymity about a topic that they wanted to disclose in order to get a trigger of self-disclosure (Ashktorab et al. 2017).

Second, suggesting avatar actions in input forms of chat systems may also facilitate effective social support during bullying experience conversations because avatar actions helped demonstrate emotional empathy, specifically, actions to express painful emotions for self-disclosure and empathy for social support. These responses to self-disclosure establish good relationships (Dai et al. 2016).

Third is facilitating to construct close relationships for providing and receiving social support. In Pigg Party, social support was effective when bullied users disclosed themselves to acquaintances. Further, in Second Life communities, anonymous users who constructed close relationships with each other provided and received emotional social support (Green-Hamann, Campbell Eichhorn, and Sherblom 2011). Constructing close friendships may decrease cyberbullying risks in self-disclosure because users who have

close friends have no occasion to disclose themselves to strangers.

## Limitations

We acknowledge that our study has the following limitations.

First, we showed the effects of social support indirectly, that is, we used usage frequencies of the avatar communication service as the measure of satisfaction of a user. Although the previous research (Rokito et al. 2019) has shown a correlation between usage frequencies of a web application and user satisfaction, conducting questionnaire surveys and/or psychological experiments should strengthen our conclusions. Additionally, sentiment analysis of words, emoticons, and emojis may reveal longer-term effects of social supports.

Second, as the bullying experience conversations in our data set are limited to conversations that contain "bullying" and its homonyms, the data set might be biased by this lexicon. Although the filtering process is essential for avoiding extensive effort for annotation, this filter drops some conversations that are related to bullying. We expect that expanding the lexicon based on the data set will mitigate this issue.

Finally, this paper has not clarified differences of effects on self-disclosure and social support between avatar communications and other communication methods, such as face-to-face, e-mail, and text communication. For clarifying this, psychological experiments to compare these communication methods about self-disclosure and social support should be conducted.

## Conclusion

We investigated the state of social support and its effectiveness between users in *Pigg Party*. We focused on users who self-disclosed their bullying experiences, and found that both verbal and nonverbal methods play important roles in facilitating social support. Avatar communication, which provides both verbal and rich nonverbal emotional expressions, contributes to effective online social support.

In general, elaborate communication methods, which are synchronous and time-consuming (e.g., face-to-face communication and avatar communication) increase closeness between people (Burke and Kraut 2014; Takano 2018). We believe that avatar communication services, providing physical-world-like communication through the Internet, are important for constructing close friendships and facilitating online social support.

## References

Aldunate, N., and González-Ibáñez, R. 2017. An Integrated Review of Emoticons in Computer-Mediated Communication. *Frontiers in Psychology* 7:2061.

Andalibi, N., and Forte, A. 2018. Announcing Pregnancy Loss on Facebook. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems - CHI '18*, 1–14. New York, New York, USA: ACM Press.

Andalibi, N.; Haimson, O. L.; De Choudhury, M.; and Forte, A. 2016. Understanding Social Media Disclosures of Sexual Abuse Through the Lenses of Support Seeking and Anonymity. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems - CHI '16*, 3906–3918. New York, New York, USA: ACM Press.

Andalibi, N.; Ozturk, P.; and Forte, A. 2017. Sensitive Self-disclosures, Responses, and Social Support on Instagram: The Case of #Depression. In *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing - CSCW '17*, 1485–1500. New York, New York, USA: ACM Press.

Antonijevic, S. 2008. FROM TEXT TO GESTURE ONLINE: A microethnographic analysis of nonverbal communication in the Second Life virtual environment. *Information, Communication & Society* 11(2):221–238.

Ashktorab, Z.; Haber, E.; Golbeck, J.; and Vitak, J. 2017. Beyond Cyberbullying: Self-Disclosure, Harm and Social Support on ASKfm. In *Proceedings of the 2017 ACM on Web Science Conference - WebSci '17*, 3–12. New York, New York, USA: ACM Press.

Blei, D. M.; Ng, A. Y.; and Jordan, M. I. 2003. Latent Dirichlet Allocation. *Journal of Machine Learning Research* 3(Jan):993–1022.

Burke, M., and Kraut, R. E. 2014. Growing closer on facebook: changes in tie strength through social network site use. In *Proceedings of the 32nd annual ACM conference on Human factors in computing systems - CHI '14*, 4187–4196. New York, New York, USA: ACM Press.

De Choudhury, M., and De, S. 2014. Mental Health Discourse on reddit: Self-Disclosure, Social Support, and Anonymity. In *Proceedings of the Eighth International Conference on Weblogs and Social Media*.

Chung, J. E. 2013. Social interaction in online support groups: Preference for online social interaction over offline social interaction. *Computers in Human Behavior* 29(4):1408–1414.

Cole, H., and Griffiths, M. D. 2007. Social Interactions in Massively Multiplayer Online Role-Playing Gamers. *CyberPsychology & Behavior* 10(4):575–583.

Dai, Y.; Shin, S. Y.; Kashian, N.; Jang, J.-w.; and Walther, J. B. 2016. The Influence of Responses to Self-Disclosure on Liking in Computer-Mediated Communication. *Journal of Language and Social Psychology* 35(4):394–411.

Ernala, S. K.; Labetoulle, T.; Bane, F.; Birnbaum, M. L.; Rizvi, A. F.; Kane, J. M.; and De Choudhury, M. 2018. Characterizing Audience Engagement and Assessing Its Impact on Social Media Disclosures of Mental Illnesses. *Twelfth International AAAI Conference on Web and Social Media (ICWSM '18)*.

Ervin-Tripp, S. M., and Lampert, M. 2009. The occasioning of self-disclosure humor. In Norrick, N. R., and Chiaro, D., eds., *Humor in Interaction*. John Benjamins Publishing Company. 3–28.

Green-Hamann, S.; Campbell Eichhorn, K.; and Sherblom, J. C. 2011. An Exploration of Why People Participate in Second Life Social Support Groups. *Journal of Computer-Mediated Communication* 16(4):465–491.

Inagaki, T. K.; Bryne Haltom, K. E.; Suzuki, S.; Jevtic, I.; Hornstein, E.; Bower, J. E.; and Eisenberger, N. I. 2016. The Neurobiology of Giving Versus Receiving Support. *Psychosomatic Medicine* 78(4):443–453.

Jaidka, K.; Guntuku, S. C.; and Ungar, L. H. 2018. Facebook versus Twitter: Differences in Self-Disclosure and Trait Prediction. *Twelfth International AAAI Conference on Web and Social Media (ICWSM '18).*

Kaczmarek, L. D., and Drążkowski, D. 2014. MMORPG Escapism Predicts Decreased Well-Being: Examination of Gaming Time, Game Realism Beliefs, and Online Social Support for Offline Problems. *Cyberpsychology, Behavior, and Social Networking* 17(5):298–302.

Kang, R.; Dabbish, L. A.; and Sutton, K. 2016. Strangers on Your Phone: Why People Use Anonymous Communication Applications. In *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing - CSCW '16*, 358–369. New York, New York, USA: ACM Press.

Koyama, N.; Tanaka, K.; Ogawa, K.; and Ishiguro, H. 2017. Emotional or Social? In *Proceedings of the 5th International Conference on Human Agent Interaction - HAI '17*, 203–211. New York, New York, USA: ACM Press.

Kralj Novak, P.; Smailović, J.; Sluban, B.; and Mozetič, I. 2015. Sentiment of Emojis. *PLOS ONE* 10(12):e0144296.

Kudo, T.; Kudo, T.; Yamamoto, K.; and Matsumoto, Y. 2004. Applying conditional random fields to Japanese morphological analysis. *Proceedings of the 2004 Conference on Empirical Methods in Natural Language Processing* 230–237.

Küster, D.; Krumhuber, E.; and Kappas, A. 2015. Nonverbal Behavior Online: A Focus on Interactions with and via Artificial Agents and Avatars. In *The Social Psychology of Nonverbal Communication*. London: Palgrave Macmillan UK. 272–302.

LaCrosse, M. B. 1975. Nonverbal behavior and perceived counselor attractiveness and persuasiveness. *Journal of Counseling Psychology* 22(6):563–566.

Liu, C.-Y., and Yu, C.-P. 2013. Can Facebook Use Induce Well-Being? *Cyberpsychology, Behavior, and Social Networking* 16(9):674–678.

Manusov, V. L., and Patterson, M. L. 2006. *The SAGE handbook of nonverbal communication*. Thousand Oaks, CA: SAGE Publications.

McKenna, K. Y. A., and Bargh, J. A. 1998. Coming out in the age of the Internet: Identity "demarginalization" through virtual group participation. *Journal of Personality and Social Psychology* 75(3):681–694.

Mehrabian, A. 1970. *Nonverbal Communication*. Chicago: Aldine-Atherton.

Ministry of Internal Affairs and Communications, Japan 2017. Present and Future of Smartphone Economy. In *2017 White Paper on Information and Communications in Japan*. Nikkei Pringing. chapter 1.

Morelli, S. A.; Lee, I. A.; Arnn, M. E.; and Zaki, J. 2015.

Emotional and instrumental support provision interact to predict well-being. *Emotion* 15(4):484–493.

Nolan, S.; Hendricks, J.; and Towell, A. 2015. Social networking sites (SNS); exploring their uses and associated value for adolescent mothers in Western Australia in terms of social support provision and building social capital. *Midwifery* 31(9):912–919.

O'Connor, E. L.; Longman, H.; White, K. M.; and Obst, P. L. 2015. Sense of Community, Social Identity and Social Support Among Players of Massively Multiplayer Online Games (MMOGs): A Qualitative Analysis. *Journal of Community & Applied Social Psychology* 25(6):459–473.

Omarzu, J. 2000. A Disclosure Decision Model: Determining How and When Individuals Will Self-Disclose. *Personality and Social Psychology Review* 4(2):174–185.

Rokito, S.; Choi, Y. H.; Taylor, S. H.; and Bazarova, N. N. 2019. Over-gratified, under-gratified, or just right? Applying the gratification discrepancy approach to investigate recurrent Facebook use. *Computers in Human Behavior* 93:76–83.

Sharma, E., and De Choudhury, M. 2018. Mental Health Support and its Relationship to Linguistic Accommodation in Online Communities. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems - CHI '18*, 1–13. New York, New York, USA: ACM Press.

Takano, M., and Mizuno, K. 2018. A Considaration of Social Relationships and Social Support in a Virtual Society. In *Proceedings of the Annual Conference of The Japanese Society for Artificial Intelligence* (in Japanese; the user interview report by the Pigg Party application provider).

Takano, M. 2018. Two types of social grooming methods depending on the trade-off between the number and strength of social relationships. *Royal Society Open Science* 5(8):180148.

The Japan Times 2017. Nagano counselors trial online consultations for troubled teens via Line messaging app.

Turner, R. J.; Turner, J. B.; and Hale, W. B. 2014. Social Relationships and Social Support. In Johnson, R. J.; Turner, R. J.; and Link, B. G., eds., *Sociology of Mental Health*. Springer Nature. 1–20.

Vlahovic, T. A.; Roberts, S.; and Dunbar, R. 2012. Effects of duration and laughter on subjective happiness within different modes of communication. *Journal of Computer-Mediated Communication* 17(4):436–450.

Wang, Y.; Tang, J.; Li, J.; Li, B.; Wan, Y.; Mellina, C.; O'Hare, N.; and Chang, Y. 2017. Understanding and Discovering Deliberate Self-harm Content in Social Media. In *Proceedings of the 26th International Conference on World Wide Web - WWW '17*, 93–102. New York, New York, USA: ACM Press.