Happiness and Sadness in Adolescents' Instagram Direct Messaging: A neural topic modeling approach

Verbeij, T.; Beyens, I.; Trilling, D. ; Valkenburg, P.M.

DOI
10.1177/20563051241229655

Publication date
2024

Document Version
Final published version

Published in
Social Media + Society

SM+S
social media + society

Article

Social Media + Society
January-March 2024: 1–13

DOI: 10.1177/20563051241229655
journals.sagepub.com/home/sms
Sage

Happiness and Sadness in Adolescents' Instagram Direct Messaging: A Neural Topic Modeling Approach

Tim Verbeij, Ine Beyens, Damian Trilling, and Patti M. Valkenburg

Abstract

We investigated the expressions of happiness and sadness in adolescents' direct messages (DMs) on Instagram. Using neural topic modeling (*BERTopic*), we analyzed 211,778 DMs belonging to 96 adolescents, who donated data from 101 Instagram accounts. Results showed that (1) expressions of happiness were more than four times more prevalent than expressions of sadness; (2) the number of DMs containing expressions of happiness and expressions of sadness were highly correlated; (3) there are temporal trends in the expression of happiness and sadness in adolescents' DMs, and there are individual differences in these trends; and (4) there is no significant between- or within-person relationship between the number of DMs containing expressions of happiness and sadness and adolescents' well-being.

Keywords

social media content, direct messages, happiness, sadness, data donation, neural topic modeling

There is much scientific and public debate about the effects of social media on adolescents' psychosocial functioning (Orben, 2020; Valkenburg et al., 2022). To date, most social media effect studies have used adolescents' *time spent* on social media as the main indicator of their social media use. However, self-reported and digital trace measures of time spent on social media cannot offer insights into the *content* of adolescents' social media use, that is, what adolescents *see* and *share* on social media. To arrive at a true understanding of the effects of social media use on psychosocial functioning, researchers need to adopt methodologies that also capture the content of adolescents' social media use (Parry et al., 2022; Valkenburg, 2022; Verbeij et al., 2022).

A promising methodology that allows researchers to obtain insights into adolescents' social media content is social media data download packages (DDPs) (Araujo et al., 2022; van Driel et al., 2022). DDPs are the personal archives of social media users that contain all their social media data stored by the platform in question. Social media users can download their personal archives from a certain social media platform and donate it to the researchers. Recent analysis of Instagram DDPs showed that sending direct messages (DMs) is one of the most popular activities among adolescents (van Driel et al., 2022). DMs allow adolescents to interact with their peers in a private and synchronous way. Yet, much remains unknown about the content that adolescents share in their DMs. Using Instagram DDPs, this study will explore the content of the DMs that adolescents share using an advanced neural topic modeling technique.

To our knowledge, eight earlier studies have analyzed the content of social media DMs, with diverging aims. Four studies investigated the linguistic characteristics of adolescents' DMs on both WhatsApp and Facebook (Hilte et al., 2021; Hilte et al., 2018, 2020; Surkyn et al., 2021). These studies showed that, for example, girls and boys adopt their conversation partner's writing style in mixed-gender conversations in DMs (Hilte et al., 2023). Another study asked adolescents and young adults to label their own DMs as either safe (e.g., about food) or unsafe (e.g., unwanted sexual solicitations). The study showed that unsafe DMs were more emotionally charged and more provocative (Ali et al., 2022). In addition, finally, three studies provided a thematic analysis of DMs on various social media platforms (Dong et al., 2006; Koch et al., 2022; Ranney et al., 2020). One of these studies showed that conflicts in adolescents' DMs could be

University of Amsterdam, the Netherlands

Corresponding Author:
Tim Verbeij, Amsterdam School of Communication Research (ASCoR), University of Amsterdam, P.O. Box 15791, 1001 NG Amsterdam, the Netherlands.
Email: t.verbeij@uva.nl

categorized as either direct versus indirect or serious versus nonserious (Ranney et al., 2020). Another study found that DMs could be categorized into six topics: sports, pornography, games, travel, entertainment, and others (Dong et al., 2006). The last study showed that certain topics are shared more by women than by men, and vice versa. For example, women talked more about social activities in their DMs (Koch et al., 2022).

Although earlier literature provided important insights into different aspects of DMs, it leaves three important gaps that this study aims to fill. A first gap is that it remains unclear how many DMs contain expressions of happiness (e.g., DMs containing positive emotions indicated by words such as "haha", "funny", or "love") or sadness (e.g., DMs containing negative emotions indicated by words such as "sad", "cry", or "regret"). To date, numerous studies have investigated expressions of happiness and sadness in public posts on social media (e.g., Larsen et al., 2015; Ofoghi et al., 2016). For example, an analysis of more than 1 billion Twitter posts showed that these posts more often contain expressions of happiness than sadness (Larsen et al., 2015). However, it is possible that this *positivity bias* in public posts on social media is less apparent in DMs because such messages are generally more intimate and allow adolescents to be their authentic self (Waterloo et al., 2018). To our knowledge, no study has focused on expressions of happiness and sadness in adolescents' DMs. Although three studies (Dong et al., 2006; Koch et al., 2022; Ranney et al., 2020) did investigate themes in DMs, they did not look at the expression of happiness or sadness. Therefore, the first aim of this study is to investigate how many of adolescents' DMs contain expressions of happiness and sadness (Research Question [RQ] 1).

As a second gap in the literature, no study so far has investigated whether and to what extent expressions of happiness and sadness in adolescents' DMs change over time. Previous studies concerning public social media posts showed temporal trends in the number of posts containing expressions of happiness and sadness (Golder & Macy, 2011; Larsen et al., 2015). For example, Golder and Macy (2011) found that public Twitter posts are more likely to contain expressions of happiness on weekends than on weekdays. On a seasonal level, they found that public posts were more likely to contain expressions of happiness when the summer solstice approaches. However, as far as we are aware, no study has determined temporal trends in the number of DMs containing expressions of happiness or sadness. Therefore, our second aim is to study whether and to what extent the number of DMs containing expressions of happiness or sadness changes over time (RQ2).

A third gap in the literature is that it is still unknown whether expressions of happiness and sadness in adolescents' DMs are related to their well-being. By now, hundreds of empirical studies have aimed to establish the relationship between adolescents' time spent on social media and their well-being, often with inconsistent results (Orben, 2020; Valkenburg et al., 2022). These mixed findings could, at least partly, be explained by the fact that time-based measures of social media use are suboptimal to investigate the relationship between social media use and well-being (Valkenburg et al., 2022). Instead of the time adolescents spend on social media, it is likely that expressions of happiness and sadness in adolescents' DMs drive the influence on well-being. Therefore, the third aim of this study is to explore the extent to which the number of DMs containing expressions of happiness or sadness is related to adolescents' well-being (RQ3).

This Study

To address the three aims of this study and the methodological limitations of previous research, we used an advanced neural topic modeling technique called *BERTopic* (Grootendorst, 2022). We analyzed 101 Instagram DDPs belonging to 96 adolescents. These Instagram DDPs contained all 211,778 DM adolescents sent or received within an 8-month period. To establish the relationship between the number of DMs containing expressions of happiness and sadness and adolescents' well-being, we focused on affective well-being, defined as adolescents' feelings of happiness (Tov, 2018). Adolescents' affective well-being was measured 12 times in a separate biweekly survey among the same group of adolescents.

We investigated the associations of expressions of happiness and sadness with well-being on a *between-person* and *within-person* level. On a *between-person* level, we investigated whether adolescents who sent or received more DMs containing expressions of happiness or sadness across the 8-month period than their peers also had higher (or lower) well-being than their peers across the same period. On a within-person level, we investigated whether adolescents had higher (or lower) well-being than they typically had during months when they sent or received more DMs containing expressions of happiness (or sadness). Earlier studies on social media use and well-being have predominately focused on between-person associations (Schønning et al., 2020). However, there is ever more consensus that investigating within-person associations is a more rigorous test of media effects because a media effect occurs within persons, reflecting an intra-individual change in cognitions, affect, or behavior of an individual due to this individual's media use (Valkenburg & Peter, 2013).

Method

This study is part of a large intensive longitudinal cohort study (https://osf.io/uxnm8) on adolescents' social media use and psychosocial functioning, which ran from 21 November 2019 to 1 July 2020. This project consisted of two 3-week Experience Sampling Methodology (ESM) studies and 16 biweekly surveys. At the end of the project, all remaining adolescents were asked if they wanted to donate their Instagram DDPs, which 102 adolescents did. For more information about DDPs, see the works by van Driel et al. (2022) and Araujo et al. (2022). This study used the DMs from the DDPs and well-being data from 12 biweekly surveys.

Participants

Participants were recruited via a large secondary school in the Netherlands. Of the 388 participants who started the overall project, 102 participants provided 110 useable Instagram DDPs. There were no age differences between participants who donated a DDP (M=14.04, SD=.69) and those who did not (M=14.14, SD=.70), $t(386)=-1.2$, $p=.23$. However, girls were more likely to donate a DDP than boys, 68% versus 52%, χ^2 (2, N=388)=14.03, p=.001, and participants from a lower educational track were less likely to donate their DDPs than those from higher educational tracks, 31% versus 48%, χ^2 (2, N=388)=9.23, p=.01. Of the 110 usable Instagram DDPs, 101 contained DMs belonging to 96 participants (M_{age}=14.01 years, SD_{age}=.67, 67% girls), of whom 95 participants (99%) identified themselves as Dutch. The educational levels of the sample were representative of the specific region in the Netherlands: 34% of participants were enrolled in the lower prevocational education track, 32% in the intermediate general education track, and 33% in the academic preparatory track.

Procedure

All participants who provided their DDPs obtained parental consent and provided informed assent. Before participants donated their DDPs, we provided them information about: (1) what they would be sharing with the research team if they agreed to participate, (2) the pseudonymization process, (3) how the data would be stored, and (4) the type of information the researchers was interested in. After participants provided assent, they automatically proceeded to an online survey in *Qualtrics*. As a first step in the survey, participants received questions about how many and what type of Instagram accounts they wanted to share (e.g., if the account was a regular account or fan account). As a second step, participants received detailed visual instructions on how to download their own DDP via the Instagram website and the smartphone app.

The actual DDP procedure consisted of two main steps. At a first step, participants had to log into their Instagram accounts to request the download and to confirm the e-mail address at which they wanted to receive the data. At a second step, within a maximum of 48 hrs after participants' request, Instagram sent the data download zip files to the e-mail address participants had provided. Once the DDPs arrived, participants needed to download the zip files from their e-mail and upload them to a protected server of the university via a private link. Each participant received their own individualized link to a data folder that was prelabeled with their participant number. Adolescents received 5 euros for each Instagram DDP they donated. A more complete description of the procedure alongside the promises and pitfalls of collecting Instagram DDPs is described in the work by van Driel et al. (2022).

Measures

Instagram DMs. Of the 110 Instagram DDPs, 101 contained the file in which the DMs were stored. This file contained both text messages and pictures or videos that were shared in the DMs. For the purpose of this study, we only extracted the text messages. In total, 211,778 text messages were shared in 3,191 unique chat sessions. The number of DMs differed per DDP and ranged from 2 to 39,141 DMs across the 8-month study period (M=2,097.81, SD=4,675.25; Med=529). The average length of a DM was 21.79 characters (SD=37.65, min=1, max=1,069).

Well-Being. Adolescents' well-being was measured in 12 biweekly surveys with the following question: "How happy did you feel in the past 7 days?" using a 7-point scale ranging from 1 (*not at all*) to 7 (*completely*), with 4 (*a little*) as the midpoint.

Statistical Analysis

We used *Python* (version 3.9.0) for all our analyses. The analysis scripts are available online on the Open Science Framework (OSF; https://osf.io/w6e9d). To determine the main topics in adolescents' Instagram DMs, we employed *BERTopic* (version 0.1.2), a neural topic modeling technique. Most previous studies have relied on *dictionary* approaches to analyze expressions of happiness and sadness in social media content, like sentiment analyses (e.g., Larsen et al., 2015). Only very few studies have employed deep-learning applications like *BERTopic*.

There are two important advantages of applications like *BERTopic* in comparison with traditional dictionary approaches. First, dictionary approaches typically rely on predefined lists of words in a certain language to determine their meaning (e.g., "happy" and "cry"). *BERTopic*, in contrast, determines the meaning of words by considering their *context*, that is, the order of and relationships between the words. For example, when an adolescent shares/receives a DM containing the word "haha", *BERTopic* also considers all the other words in this DM. A second related advantage of *BERTopic* is that certain unknown or slang words, as common in adolescents' language, which would not show up in predefined lists, can also be recognized and clustered with other expressions of happiness and sadness. For example, while a word like "omg" is not part of Dutch language, it is often used among Dutch adolescents. In *BERTopic*, such a word would cluster together with known expressions of happiness, such as "funny" or "haha". This focus on adolescents' language is important because neural network models have mainly been trained to predict expressions of happiness or sadness in *adults*' social media content (e.g., Ouyang et al., 2015; Severyn & Moschitti, 2015), but have hardly been used to assess expressions of happiness and sadness in *adolescents'* social media content.

In all, neural network models, like *BERTopic*, provide us with a richer understanding of adolescents' language in their Instagram DMs compared with traditional dictionary approaches. Yet, even neural network models cannot solve the *symbol grounding problem* in artificial intelligence (AI), that is, how neural network models connect words (symbols) to the real-world objects they refer to. Therefore, it remains a challenge for these models to truly understand the meaning of words (for a discussion, see the work by Harnad, 1990). However, *BERTopic* does consider other words in adolescents' Instagram DMs to determine the "meaning" of words, which makes it more versatile in interpreting language compared with dictionary approaches.

As discussed by Grootendorst (2022), *BERTopic* generates topics through three main steps: (1) generating DM embeddings, (2) clustering the DMs, and (3) generating keywords for the topics. In a first step, DMs are embedded to create representations in a vector space. This allows researchers to compare the semantic meaning of the DMs. The closer adolescents' DMs are in the vector space, the more semantically similar they are (e.g., "happy" and "funny"). Conversely, the further they are from each other in this vector space, the more semantically unique they are (e.g., "happy" and "sad"). Compared with more traditional approaches toward topic modeling (e.g., Latent Dirichlet Allocation [LDA]), embedding-based approaches like *BERTopic* show superior performance (Grootendorst, 2022). While there are multiple embedding-based models available (e.g., top2vec, see the work by Angelov, 2020), we decided to use *BERTopic* because it has been shown to perform well compared with many alternatives (Grootendorst, 2022).

In a second step, the embeddings are clustered using a clustering technique called Hierarchical Density-Based Spatial Clustering of Applications with Noise (*HDBSCAN*; for more details about *HDBSCAN*, see the work by McInnes et al., 2017). This clustering technique groups DMs that are similar to each other. *HDBSCAN* also allows DMs that do not belong to any other topic to be modeled as outliers (i.e., "Topic-1"). Other unsupervised techniques for identifying topics, such as LDA, lack the advantage of approaches that employ contextualized embeddings to distinguish homonyms or deal with (near-)synonyms. Neither do they allow outliers to be modeled as a separate topic (Li et al., 2018). This makes these topic modeling techniques more sensitive to noise that occurs frequently in short-text documents like DMs. In the third and final step, *BERTopic* generates keywords for each topic, using term frequency–inverse document frequency (TF–IDF) scores. The TF–IDF score shows which words are most important in a DM by weighing the frequency in a DM by how rare the words are in other DMs (for more details, see the work by Manning et al., 2008). These TF–IDF scores also allow researchers to reduce the number of topics by merging topics that have similar keywords (e.g., topics with the keywords "funny" and "laugh").

Topic Models. We compared nine different *BERTopic* models, which varied in topic size and message length. In each model, we defined the minimum size of each topic through the parameter *min_topic_size*. The lower this parameter is set, the smaller topics are allowed to be (i.e., the fewer DMs a topic contains). This also allows niche topics (i.e., topics in DMs that are not discussed often). We compared models with a minimum topic size of 50, 100, and 200 DMs. In addition, we explored variations in DM length, because short messages like DMs have been proven challenging for topic modeling techniques (Mehrotra et al., 2013). Based on the work by Mehrotra et al. (2013), we compared models based on three different DM lengths: the original DMs (i.e., with at least one character), DMs containing at least 20 characters, and concatenating the DMs per conversation per hour.

Since multiple languages (e.g., Dutch and English) were used in the DMs, we employed a multilingual model ("*paraphrase-multilingual-MiniLM-L12-v2*"), which also deals well with emojis. That is, *BERTopic* treats emojis just like words and embeds them in the same space, such that the emoji's meaning is captured. After fitting each topic model, we used the *CountVectorizer* function of the package *Scikit-learn* (Pedregosa et al., 2011) to preprocess the topic model to avoid that stop words (e.g., "while" or "because") would be used in the generation of keywords. In this case, a *CountVectorizer* counts how many times each stop word appears in adolescents' DMs. This is a suggested method by the creator of *BERTopic* (Grootendorst, 2022), especially useful for short DMs, as they typically contain many stop words. It is important to note that we explicitly fitted this vectorizer *after* we fitted our model so it would not actually change the meaning of the DMs.

Final Topic Model. All nine models identified similar trends in topics, but the models varied in terms of number of topics (from 13 to 669 topics) and face validity (i.e., the interpretability of the topics). The best model in terms of face validity was the model with a minimum topic size of 200 DMs and a DM length of at least 20 characters. This model contained 77,142 DMs of 95 participants and identified 60 topics. This was considered the best model for two reasons. First, *BERTopic* needs Instagram DMs that are sufficiently long so that the meaning of words can be determined by its context. This is more difficult with very short Instagram DMs (e.g., <20 characters) as they are often used without context. Second, we cannot assume that Instagram DMs concatenated within a conversation across a specific time interval are about the same topic. In other words, the best time interval for concatenation is different for every participant and every conversation. For instance, on 1 day, two people might talk intensively sending Instagram DMs to each other within minutes, while at other moments the same two people might wait days before replying to an Instagram DM. As is common in topic modeling, we determined topics based on the top-10 keywords (e.g., Porturas & Taylor, 2021; Song et al., 2009). As a rule, topics were considered to contain expressions of happiness if they were determined by keywords such as "nice", "hahaha", "omg", "funny", and "beautiful", and expressions of sadness if they were determined by keywords such as "dead", "cry", "regret", "feel", and "sad". We manually merged three topics of which the keywords indicated expressions of happiness and three topics that showed expressions of sadness.

Validation of the Final Topic Model. We concentrated our validation efforts on whether the topics that contained expressions of happiness and sadness could be reliably labeled as such by two annotators as well as the final model. We took a stratified subsample of 60 DMs, of which 30 were labeled by the final model as containing expressions of happiness and 30 as containing expressions of sadness. Two researchers independently coded the 60 DMs as either happiness or sadness expressions. The interrater reliability (Cohen's κ) between the researchers and between the researchers and the predictions made from our final model ranged from substantial to near-perfect agreement ($\kappa_{rater.a*rater.b}=.83$; $\kappa_{rater.a*model}=.77$; and $\kappa_{rater.b*model}=.83$).

To further validate our model, we compared the results of the topic model with the results of manual coding of the DMs that also took into account the context of the DMs (van Atteveldt et al., 2021). That is, for our stratified sample of 60 Instagram DMs, we also considered the DMs that preceded the respective DM in the chat session to determine whether a DM contained expressions of happiness or sadness, respectively. Since there is no optimal way to concatenate conversations, this type of topic determination is difficult to realize with *BERTopic* modeling. The confusion matrix of the manual coding of the 60 DMs and the predictions made by our final *BERTopic* model are available on OSF (https://osf.io/q5gz4).

We also calculated precision, that is, the percentage of model predictions that were correct. A precision value of 1.0 indicates that there are no false positives. The precision was .93 for the DMs containing expressions of happiness and 1.00 for the DMs containing expressions of sadness. We also calculated recall, the percentage of manually coded happiness/sadness expressions that were also identified by the model as such. A recall value of 1.0 indicates that there are no false negatives. The recall was 1.00 for the DMs containing expressions of happiness and .94 for the DMs containing expressions of sadness. To conclude, our final model can, in nearly all instances, correctly detect DMs that contain expressions of happiness or sadness and distinguish between them.

Investigating RQs. RQ1 asked about the frequency of DMs containing expressions of happiness or sadness. To investigate RQ1, we calculated the total frequency of DMs that were labeled as containing expressions of happiness or sadness across the sample (RQ1a). In addition, for each adolescent, we calculated the ratio of happiness versus sadness expressions in their Instagram DMs (RQ1b). RQ2 asked whether the frequency of DMs that contained expressions of happiness or sadness changed over time. To investigate RQ2, we employed the dynamic topic modeling function of *BERTopic* to visualize how the number of DMs containing expressions of happiness or sadness changed over the course of the study period across all adolescents (RQ2a) and for each adolescent separately (RQ2b).

RQ3 asked about the association between the frequency of DMs labeled as containing expressions of happiness or sadness and well-being. We assessed this relationship on both a between- and within-person level. On the *between-person* level, we investigated whether adolescents who sent or received more DMs containing expressions of happiness or sadness than their peers also had a higher (or lower) level of well-being than their peers (RQ3a). To determine the between-person *Pearson* correlation coefficient, we used the *corr* function of the Python package *pingouin* (Vallat, 2018). We interpreted the between-person correlations according to the guidelines of Gignac and Szodorai (2016). A correlation ranging from (1) $-.10<r<.10$ was interpreted as "nonexistent to very small"; (2) $.10\leq r<.20$ was interpreted as "small"; (3) $.20\leq r<.30$ was interpreted as "moderate"; and (4) $r\geq .30$ was interpreted as "large".

Finally, on the *within-person* level, we investigated whether adolescents had higher (or lower) well-being than they typically had during months when they sent or received more DMs containing expressions of happiness (or sadness) (RQ3b). To get indices of well-being on a *monthly* time interval, we calculated average well-being scores across the two biweekly surveys of each month. To calculate the within-person associations, we used the *rm_corr* function of the Python package *pingouin* (Vallat, 2018). Following Meier and Reinecke (2021), we interpreted the within-person correlations ranging from −.05 to +.05 as "nonexistent to very small", and within-person correlations beyond this range as either negative or positive.

Results

Topics and Themes in Adolescents' DMs

Our final *BERTopic* model consisted of 60 topics. An overview of the 20 largest topics and their corresponding keywords is presented in Appendix A (https://osf.io/nrjc6). Figure 1 shows a visualization of the 60 topics projected into a two-dimensional vector space. The closer the topics are in the vector space, the more similar they are, and the further apart they are, the more dissimilar they are. For example, as

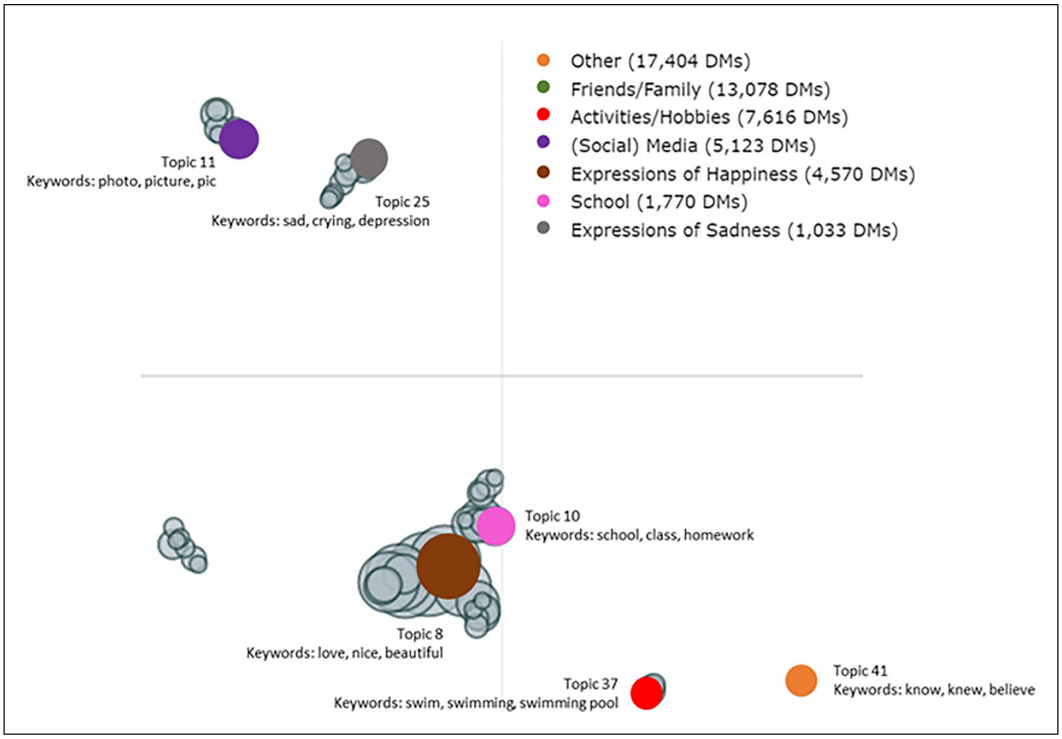

Figure 1. Representation of the final *BERTopic* model projected into a two-dimensional space.
Note. Each circle represents a topic. Larger circles are indicative of larger topic sizes (i.e., more DMs belong to that topic). The closer topics are in the vector space, the more similar they are.

Figure 1 shows, the two topics covering expressions of happiness and sadness are in the opposite dimension of the vector space.

We manually merged three topics of which the keywords indicated expressions of happiness and three topics that showed expressions of sadness. A total of 5,603 DMs (7%) were labeled as containing expressions of either happiness or sadness. A total of 27,587 DMs (36%) that did not contain an expression of happiness or sadness were manually clustered into four other themes based on overlap in their keywords. These themes were friends/family (17%), activities/hobbies (10%), (social) media (7%), and school (2%). A total of 17,404 DMs (23%) did not belong to one of these themes and were manually clustered together as a separate theme ("other"). A total of 26,548 DMs (34%) could not be labeled at all and were considered outliers. The number of outliers was consistent across adolescents ($M=35\%$, $SD=13\%$). Figure 2 provides a visualization of the DMs and their respective categorization. Since our study is focused on well-being, in the remainder of this article, we will only focus on expressions of happiness and sadness in adolescents' DMs.

Expression of Happiness and Sadness in Adolescents' DMs

To address RQ1, we investigated the frequency of DMs containing expressions of happiness or sadness. Overall (RQ1a), expressions of happiness exceeded expressions of sadness: Of the 77,142 DMs, 4,570 (6%) were labeled as containing expressions of happiness, while 1,033 DMs (1%) were labeled as containing expressions of sadness. This preponderance of happiness over sadness in the sample was also observed on an individual level (RQ1b). Of the 95 adolescents in our final model, 74 (78%) sent or received more DMs that contained expressions of happiness than sadness, 20 adolescents (21%) sent or received an equal number of DMs that had expressions of happiness and sadness, and one adolescent (1%) sent or received more DMs that had expressions of sadness than happiness. To conclude, both across the sample and on an individual level, DMs that contained expressions of happiness were more prevalent than DMs that contained expressions of sadness.

Table 1 presents the correlations between expressions of happiness and sadness. Overall, we found that, on a between-person and within-person level, the number of DMs containing expressions of happiness and sadness were highly correlated. The between-person correlation ($r=.95$) indicates that adolescents who sent or received more DMs containing expressions of happiness compared with their peers also sent or received more DMs containing expressions of sadness. On a within-person level, we found a similar trend ($r=.76$). That is, when adolescents sent or received more DMs containing expressions of happiness than they usually did, they also sent or received more DMs containing expressions of sadness. The intraclass correlations (ICCs) indicated that 68% of the variance in sending/receiving expressions of happiness and

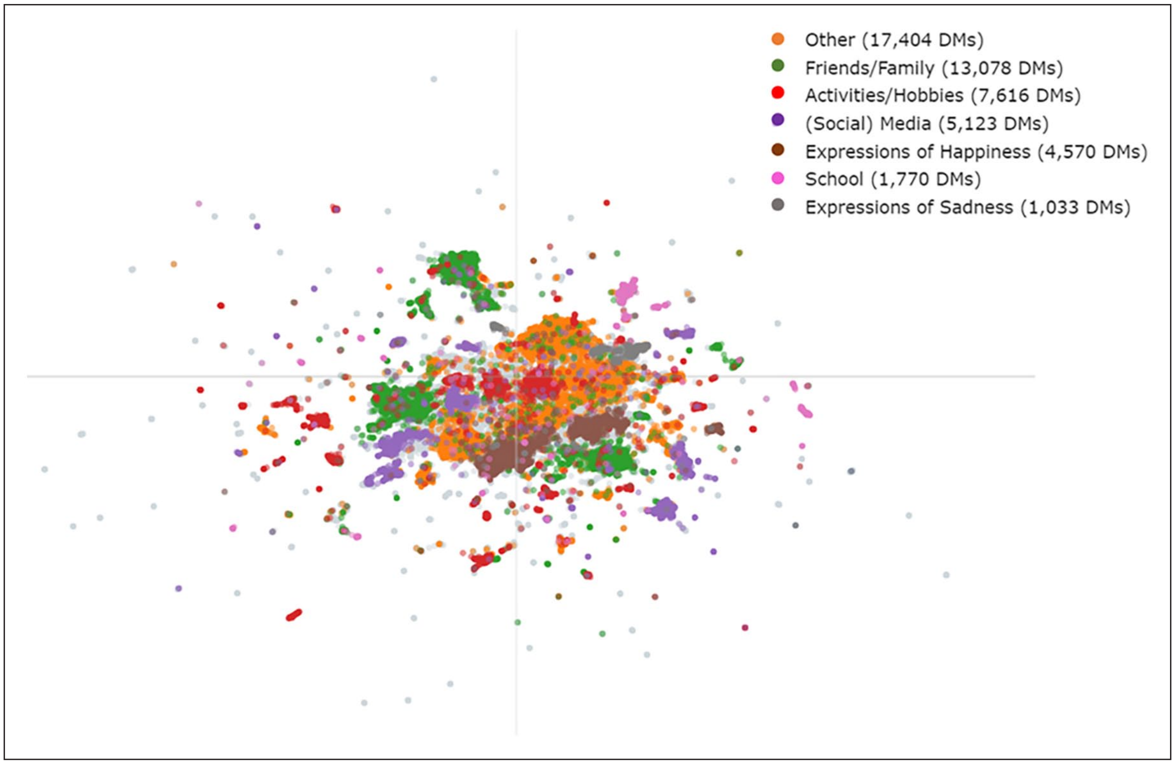

Figure 2. Categorization of the 77,142 direct messages into six themes projected into a two-dimensional vector space.
Note. The dots represent adolescents' DMs. Light gray dots represent outliers. The closer DMs are in the vector space, the more similar they are.

Table 1. Descriptive Statistics and Between-Person, Within-Person, and Intraclass Correlations for All Variables.

	EH	ES	WB
EH	—	.76*	−.04
ES	.95*	—	−.04
WB	−.07	−.09	—
M	48.11	10.87	5.61
SD	138.11	34.45	1.14
Range	0–1,227	0–301	1–7
ICC	.68	.61	.55

Note. Correlations below the diagonal line represent between-person correlations; correlations above the diagonal line represent within-person correlations. EH: number of DMs containing expressions of happiness; ES: number of DMs containing expressions of sadness; ICC: intraclass correlation; WB: well-being.
*$p < .001$.

61% of the variance in sending/receiving expressions of sadness were explained by differences between adolescents, while 32% and 39% were explained by fluctuations within adolescents, respectively.

Temporal Trends in the Expression of Happiness and Sadness in Adolescents' DMs

To address the second aim of our study, we investigated whether the frequency in DMs that contained expressions of happiness and sadness changed over time. We investigated these temporal trends on both a sample level (RQ2a) and an individual level for all adolescents separately (RQ2b). We looked at temporal trends over the full 8-month study period. It should be noted that our study started on 21 November . Therefore, for November, only 10 days could be included in the analyses.

Temporal Trends Across All Adolescents. Overall, as the red line in Figure 3 shows, adolescents sent or received DMs that contained expressions of happiness least often in November, December, January, February, May, and June. Conversely, there was a peak in the number of DMs containing expressions of happiness in March and April. DMs that contained expressions of sadness (blue line in Figure 3) were least often sent in the first 4 months of the study (November until February) and most often in the final 4 months of the study (March until June). To conclude, across adolescents, the number of DMs that contained expressions of happiness or sadness was most prevalent in March and April.

Temporal Trends on an Individual Level. On an individual level, we explored temporal trends in expressions of happiness and sadness in DMs for each adolescent separately and potential differences across adolescents in these temporal trends. We found that temporal trends were different for each adolescent. Figure 4 shows the temporal trends for two participants, who differ in the number of DMs that they sent or received.

The left-hand plot shows that the number of DMs that contained expressions of sadness remained relatively constant over time for this participant. However, the participant's expression of happiness declined over time, such that the expression of happiness was as low as the expression of sadness by the end of the 8-month period. In contrast, the right-hand plot in Figure 4 shows a participant who sent or received least DMs that contained expressions of happiness and sadness in the first 4 months of the study and sent or received most of such DMs midway the study. For this person, an increase in the number of DMs that contained expressions of happiness seemed to co-occur with an increase in the number of DMs that had expressions of sadness. Altogether, Figure 4 illustrates that individual differences exist in the temporal trends in the expressions of happiness and sadness.

The Relationship between Expressions of Happiness and Sadness in DMs and Well-Being

To address the third aim of our study, we investigated the association between the frequency of DMs containing expressions of happiness or sadness and adolescents' well-being. We investigated this association on a between-person (RQ3a) and within-person level (RQ3b). On a between-person level, adolescents who sent or received more DMs containing expressions of happiness ($r=-.07$, $p=.54$, 95% confidence interval [CI]=[−.27, .14]) or sadness ($r=-.09$, $p=.40$, 95% CI=[−.29, .12]) than their peers did not have higher (or lower) well-being than their peers. On a within-person level, when

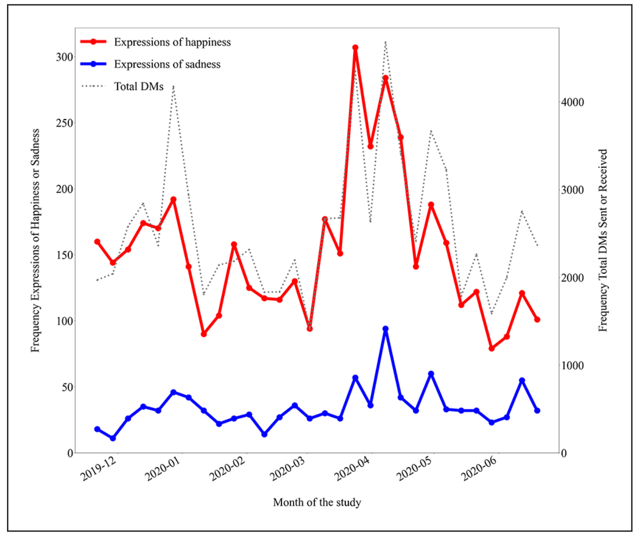

Figure 3. The number of direct messages containing expressions of happiness (red line) and sadness (blue line) over time.
Note. The gray line reflects the total number of DMs. Each dot represents a week of the study.

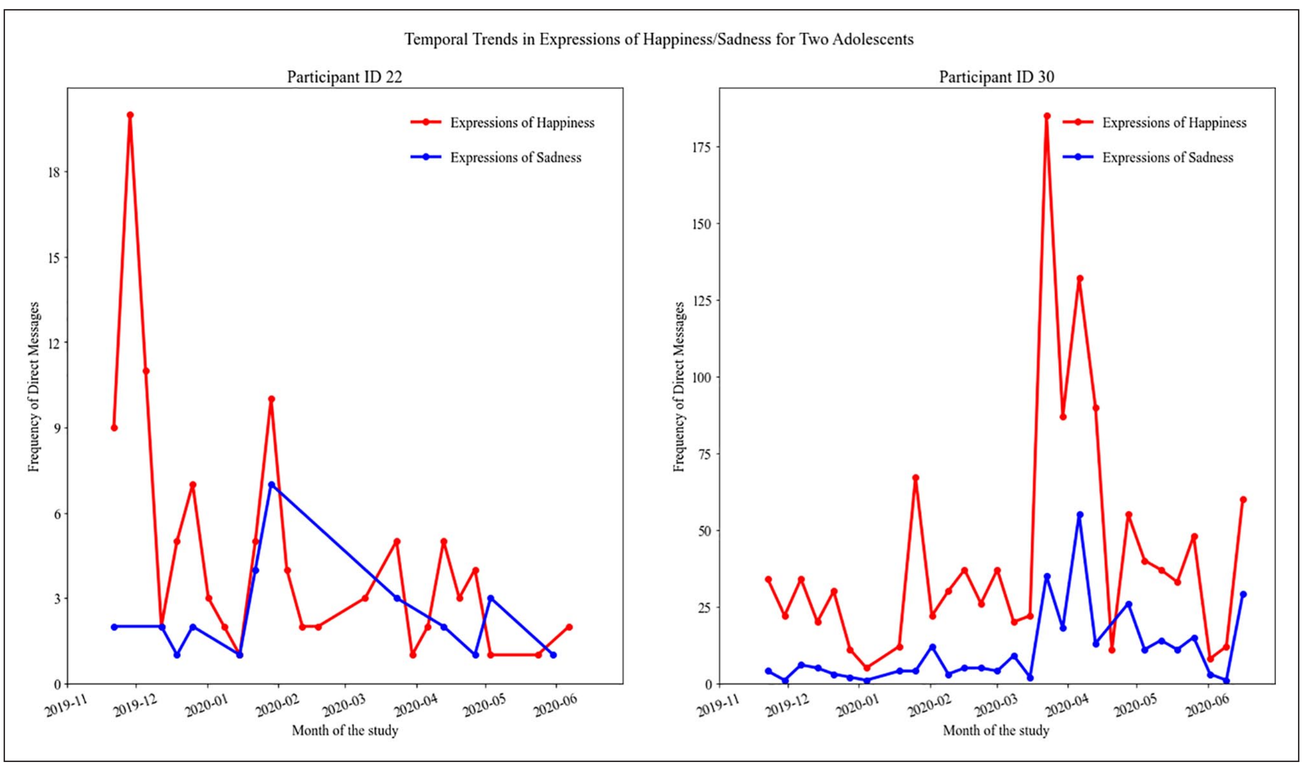

Figure 4. Temporal trends in the frequency of direct messages containing expressions of happiness (red line) and sadness (blue line) for two different participants.
Note. The *y*-axis differs per plot.

adolescents sent or received more DMs containing expressions of happiness ($r = -.04$, $p = .44$, 95% CI $= [-.14, .06]$) or sadness ($r = -.04$, $p = .46$, 95% CI $= [-.14, .06]$) than they usually did, they did not have higher (or lower) well-being than they usually did. To conclude, both on a between-person and within-person level, the frequency of DMs containing expressions of happiness or sadness was not associated with adolescents' well-being.

Sensitivity Analyses

To determine the robustness of our results, we performed two sensitivity analyses. First, since the majority (51%) of the Instagram DMs in our final topic model were sent or received by just seven Instagram accounts, we explored whether our main findings still hold when these accounts were excluded. Second, we explored the robustness of our results by distinguishing the Instagram DMs that adolescents sent and those they received. The results of these sensitivity analyses did not deviate from our main findings. More details about these sensitivity analyses are available on OSF (https://osf.io/xaz76).

Discussion

The aim of this study was to provide insights into the content of adolescents' social media use, using state-of-the-art neural topic modeling (*BERTopic*) (Grootendorst, 2022). In a sample of 96 adolescents who donated their Instagram data download packages (DDPs), we investigated the number of DMs that contained expressions of happiness and sadness and their respective relationship with well-being. The four main findings of our study can be summarized as follows: (1) expressions of happiness were sent or received four times more than expressions of sadness; (2) the number of DMs containing happiness and the number of DMs containing sadness were highly correlated; (3) there are temporal trends in the expression of happiness and sadness in adolescents' DMs, and these temporal trends were different for each adolescent; and (4) there is no significant between- or within-person relationship between the number of messages containing expressions of happiness and sadness and adolescents' well-being.

Similar to studies that investigated expressions of happiness and sadness in public social media posts using vastly different methodologies (Larsen et al., 2015), we found that adolescents sent and received more DMs that contained expressions of happiness than sadness; happiness was shared more than four times more often than sadness. These findings are also in line with studies by Valkenburg et al. (2021) and Weinstein (2018), showing that adolescents' experiences on social media are more often positive than negative. Our results confirm that there is a *positivity bias* in the expression of happiness and sadness on social media (Reinecke & Trepte, 2014). While previous studies showed that the positivity bias is apparent in public social media posts (Schreurs & Vandenbosch, 2021), our results revealed that this bias also applies to adolescents' private DMs. Yet, whereas the positivity bias in public social media posts mostly leads to negative effects, like envy, this bias in private DMs might also yield enjoyment or inspiration (Meier & Johnson, 2022; Weinstein, 2018).

Our study also showed that adolescents differ in their expression of happiness and sadness. We found that adolescents who sent or received more happy DMs compared with other adolescents, also sent or received more sad DMs. As such, it seems that some adolescents are more emotionally expressive than other adolescents, regardless of the valence of their emotions (i.e., happiness or sadness). These findings corroborate the notion that there are individual differences in the extent to which adolescents outwardly display their emotions (Kring et al., 1994). An open question is whether these individual differences are also reflected in adolescents' psychosocial functioning. For example, research concerning face-to-face interactions found that adolescents with low emotional expressivity also have poorer peer relations and perform worse at school (Zeman et al., 2006). It is likely that these findings also translate to adolescents' online interactions.

We further extended the literature by investigating temporal trends in the expression of happiness and sadness in adolescents' DMs. We found that adolescents sent or received most expressions of happiness and sadness in March and April. This is most likely due because this period coincided with the start of the coronavirus disease 2019 (COVID-19) lockdown. During this lockdown, adolescents were more-or-less forced to communicate with their peers via DMs, as real-life meetings were constrained. In June, when most COVID-19 regulations were relaxed in the Netherlands and adolescents were allowed to go back to school, we found fewer DMs containing expressions of happiness or sadness. These findings suggest that social media allow adolescents to stay in contact with their peers in extreme external circumstances (e.g., COVID-19) and express both their positive and negative emotions. When face-to-face contact is limited, social media might be a beneficial tool for sustaining adolescents' normal emotional development.

We also investigated the relationship between adolescents' expressions of happiness and sadness with adolescents' well-being. Most previous studies that have looked at the relationship between social media use and well-being have used *time spent* on social media as the main indicator of social media use. The majority of these studies reported nonsignificant relationships between social media use and well-being (for reviews, see the works by Orben, 2020 and Valkenburg et al., 2022). In line with these studies, our results showed that social media content, in terms of the number of DMs containing expressions of happiness and sadness, was not associated with adolescents' well-being on a monthly level. There might be at least two explanations for this finding. First, the effect of adolescents' expression of happiness and sadness in DMs on their well-being might differ from adolescent to adolescent (Beyens et al., 2020, 2021; Valkenburg et al., 2021). That is, some adolescents might become happier when sending/receiving DMs containing expressions of happiness or sadness, whereas others might become less happy or stay unaffected. In our study, these individual differences might have canceled each other out, resulting in the overall null effect. A second explanation might be that the monthly time interval in which we measured expressions of happiness and sadness might have been too long to validly assess associations with well-being. That is, it is possible that expressions of happiness and sadness in adolescents' DMs predict well-being on a shorter time scale, for example, within days or even hours (Beyens et al., 2020; Boele et al., 2023).

Strengths, Limitations, and Avenues for Future Research

As a first strength, our data download approach allowed us to collect the DM adolescents *sent and received* in the study period rather than only the DM adolescents sent. As such we were able to capture not only active (sending DMs) but also passive (receiving DMs) social media use. This is important because DMing cannot be dichotomized into active and passive use, as it is a dynamic process of both sending and receiving DMs (Meier & Krause, 2023). As a second strength, our intensive longitudinal design allowed us to collect Instagram DDPs and survey data on adolescents' well-being across an 8-month period. This allowed us to investigate the *within-person associations* between adolescents' DMs containing expressions of happiness and sadness and well-being. This focus on within-person processes is important, because a media effect is a within-person effect, and each adolescent might have his (or her) own unique susceptibility to the effects of social media content (Valkenburg & Peter, 2013). As a third strength, this study and its neural topic modeling approach allowed us to comprehensively capture adolescent language during a developmental phase that is characterized by many changes in linguistics (e.g., the use of slang words) and personality (e.g., emotional changes).

Despite its strengths, our study also has some limitations. First, there was a skewed distribution of the number of messages adolescents sent or received: More than half (51%) of all DMs came from just seven Instagram accounts. This skewed distribution might have affected the distribution of topics. For example, if one of the most frequent DMers sent or received many DMs containing expressions of happiness, it is likely that this topic became more pronounced in our topic model. However, our sensitivity analysis showed that excluding these seven Instagram accounts lead to a similar distribution of DMs containing expressions of happiness and sadness. Second, it is important to note that we did not focus on intensity in happiness and sadness. For instance, being annoyed and frustrated are both expressions of sadness but differ in their intensity (De Choudhury et al., 2012). As a consequence, these nuances in expressions of happiness and sadness might also influence adolescents' well-being in different ways. Third, although neural network models, like *BERTopic*, have multiple advantages over traditional dictionary approaches, they are less straightforward to understand than dictionary approaches. These neural network models are often assumed a "black box". That is, the corpus goes in (e.g., adolescents' Instagram DMs) and the topics come out, while it is unclear how the model determines the topics. However, the field of Explainable AI (XAI) has come up with multiple approaches to combat this limitation of neural network models, explaining how predictions made by such "black-box" models can be interpreted using measures of feature importance, surrogate models, visualizations, and so on (e.g., see the work by Molnar, 2022).

We have three suggestions for future research. First, researchers should aim to combine different types of content (e.g., both text and pictures/videos from DMs). In this study, we focused on the text messages in adolescents' DMs. However, adolescents' DMs also contain visual information, such as pictures and videos. To get a more comprehensive picture of the expressions of happiness and sadness specifically and content in DMs more generally, researchers should combine results from analyzing the expressions of happiness and sadness in text messages (e.g., via *BERTopic*) and analyzing expressions of happiness and sadness in pictures (e.g., using computer vision) (Hu et al., 2014). For instance, multimodal transformer models are promising as they allow researchers to analyze images and text simultaneously. Second, researchers should aim to analyze the expressions of happiness and sadness in adolescents' DMs on different social media platforms, because adolescents use different platforms in complementary ways (van Driel et al., 2019). For instance, Instagram might be primarily used by adolescents to communicate with their peers, whereas WhatsApp is used to communicate with their parents. In turn, the prevalence of expressions of happiness and sadness might differ across social media platforms. Third, it would be interesting to systematically investigate how different neural network topic modeling approaches (e.g., *top2vec* versus *BERTopic*) could lead to different results when investigating adolescents' social media content, like adolescents' Instagram DMs.

Conclusion

Through neural topic modeling, this study provided novel insights into adolescents' expressions of happiness and sadness in their DMs. Adolescents were more likely to express happiness than sadness in their DMs. We also showed that the number of messages containing expressions of happiness or sadness changes over time and that these trends differ per adolescent. Finally, we showed that there is a nonsignificant relationship between the number of messages containing happiness or sadness and adolescents' well-being. At the moment, much is still unknown about the content that adolescents share in their DMs and its relationship with their well-being. Novel methods allow for novel results; advances in AI, like neural topic modeling, might be an important avenue to help social media researchers gain a more thorough understanding of adolescents' social media content.

Acknowledgements

The authors acknowledge Irene I. van Driel and Teun Siebers for their contribution to the data collection of this study.

Author Contributions

T.V. wrote the article. I.B., D.T., and P.M.V. commented on drafts of the article. T.V. and D.T. analyzed the data.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This preregistered study was funded by an NWO Spinoza Prize and a Gravitation grant (NWO Grant 024.001.003; Consortium on Individual Development) awarded to P.M.V. by the Dutch Research Council (NWO).

ORCID iDs

Tim Verbeij https://orcid.org/0000-0003-4684-3620
Ine Beyens https://orcid.org/0000-0001-7023-867X
Damian Trilling https://orcid.org/0000-0002-2586-0352

References

Ali, S., Razi, A., Kim, S., Alsoubai, A., Gracie, J., De Choudhury, M., Wisniewski, P. J., & Stringhini, G. (2022, April 29–May 5). *Understanding the digital lives of youth: Analyzing media shared within safe versus unsafe private conversations on Instagram* [Conference session]. CHI'22: Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems, New Orleans, LA, United States.

Angelov, D. (2020). *Top2vec: Distributed representations of topics* [Preprint]. arXiv. https://doi.org/10.48550/arXiv.2008.09470

Araujo, T., Ausloos, J., van Atteveldt, W., Loecherbach, F., Moeller, J., Ohme, J., Trilling, D., van de Velde, B., de Vreese, C., & Welbers, K. (2022). OSD2F: An open-source data donation framework. *Computational Communication Research*, *4*(2), 372–387. https://doi.org/10.5117/CCR2022.2.001.ARAU

Beyens, I., Pouwels, J. L., van Driel, I. I., Keijsers, L., & Valkenburg, P. M. (2020). The effect of social media on well-being differs from adolescent to adolescent. *Scientific Reports*, *10*, Article 10763. https://doi.org/10.1038/s41598-020-67727-7

Beyens, I., Pouwels, J. L., van Driel, I. I., Keijsers, L., & Valkenburg, P. M. (2021). Social media use and adolescents' well-being: Developing a typology of person-specific effect patterns. *Communication Research*, *0*(0). https://doi.org/10.1177/00936502211038196

Boele, S., Nelemans, S. A., Denissen, J. J. A., Prinzie, P., Bülow, A., & Keijsers, L. (2023). Testing transactional processes between parental support and adolescent depressive symptoms: From a daily to a biennial timescale. *Development and Psychopathology*, *35*(4), 1656–1670. https://doi.org/https://doi.org/10.1017/S0954579422000360

De Choudhury, M., Counts, S., & Gamon, M. (2012). Not all moods are created equal! Exploring human emotional states in social media. *Proceedings of the International AAAI Conference on Web and Social Media*, *6*(1), 66–73. https://doi.org/10.1609/icwsm.v6i1.14279

Dong, H., Hui, S. C., & He, Y. (2006). Structural analysis of chat messages for topic detection. *Online Information Review*, *30*(5), 496–516. https://doi.org/10.1108/14684520610706398

Gignac, G. E., & Szodorai, E. T. (2016). Effect size guidelines for individual differences researchers. *Personality and Individual Differences*, *102*, 74–78. https://doi.org/10.1016/j.paid.2016.06.069

Golder, S. A., & Macy, M. W. (2011). Diurnal and seasonal mood vary with work, sleep, and daylength across diverse cultures. *Science*, *333*(6051), 1878–1881. https://doi.org/10.1126/science.1202775

Grootendorst, M. (2022). *BERTopic: Neural topic modeling with a class-based TF-IDF procedure* [Preprint]. arXiv. https://doi.org/10.48550/arXiv.2203.05794

Harnad, S. (1990). The symbol grounding problem. *Physica D: Nonlinear Phenomena*, *42*(1–3), 335–346.

Hilte, L., Daelemans, W., & Vandekerckhove, R. (2021). Interlocutors' age impacts teenagers' online writing style: Accommodation in intra- and intergenerational online conversations. *Frontiers in Artificial Intelligence*, *4*, Article 738278. https://doi.org/10.3389/frai.2021.738278

Hilte, L., Vandekerckhove, R., & Daelemans, W. (2018). Social media writing and social class: A correlational analysis of adolescent CMC and social background. *International Journal of Society, Culture & Language*, *6*(2), 73–89.

Hilte, L., Vandekerckhove, R., & Daelemans, W. (2023). Linguistic accommodation in teenagers' social media writing: Convergence patterns in mixed-gender conversations. *Journal of Quantitative Linguistics*, *29*(2), 241–268. https://doi.org/10.1080/09296174.2020.1807853

Hu, Y., Manikonda, L., & Kambhampati, S. (2014). What we instagram: A first analysis of instagram photo content and user types. *Proceedings of the International AAAI Conference on Web and Social Media*, *8*(1), 595–598. https://doi.org/10.1609/icwsm.v8i1.14578

Koch, T. K., Romero, P., & Stachl, C. (2022). Age and gender in language, emoji, and emoticon usage in instant messages. *Computers in Human Behavior*, *126*, Article 106990. https://doi.org/10.1016/j.chb.2021.106990

Kring, A. M., Smith, D. A., & Neale, J. M. (1994). Individual differences in dispositional expressiveness: Development and validation of the Emotional Expressivity Scale. *Journal of Personality and Social Psychology*, *66*(5), 934–949.

Larsen, M. E., Boonstra, T. W., Batterham, P. J., O'Dea, B., Paris, C., & Christensen, H. (2015). We feel: Mapping emotion on Twitter. *IEEE Journal of Biomedical and Health Informatics*, *19*(4), 1246–1252. https://doi.org/10.1109/JBHI.2015.2403839

Li, X., Wang, Y., Zhang, A., Li, C., Chi, J., & Ouyang, J. (2018). Filtering out the noise in short text topic modeling. *Information Sciences*, *456*, 83–96. https://doi.org/10.1016/j.ins.2018.04.071

Manning, C. D., Raghavan, P., & Schütze, H. (2008). *Introduction to information retrieval*. Cambridge University Press. https://doi.org/10.1017/CBO9780511809071

McInnes, L., Healy, J., & Astels, S. (2017). Hdbscan: Hierarchical density based clustering. *The Journal of Open Source Software*, *2*(11), Article 205.

Mehrotra, R., Sanner, S., Buntine, W., & Xie, L. (2013, July 28–August 1). *Improving LDA topic models for microblogs via tweet pooling and automatic labeling* [Conference session]. Proceedings of the 36th International ACM SIGIR Conference on Research and Development in Information Retrieval, Dublin, Ireland.

Meier, A., & Johnson, B. K. (2022). Social comparison and envy on social media: A critical review. *Current Opinion in Psychology*, *45*, Article 101302. https://doi.org/10.1016/j.copsyc.2022.101302

Meier, A., & Krause, H.-V. (2023). Does passive social media use harm well-being? An adversarial review. *Journal of Media Psychology*, *35*(3), 169–180. https://doi.org/10.1027/1864-1105/a000358

Meier, A., & Reinecke, L. (2021). Social media and mental health: Reviewing effects on eudaimonic well-being. In R. L. Nabi & J. G. Myrick (Eds.), *Emotions in the digital world: Exploring affective experience and expression in online interactions*. Oxford University Press.

Molnar, C. (2022). *Interpretable machine learning: A guide for making black box models explainable* (2nd ed.). https://christophm.github.io/interpretable-ml-book

Ofoghi, B., Mann, M., & Verspoor, K. (2016). Towards early discovery of salient health threats: A social media emotion classification technique. *Pacific Symposium on Biocomputing. Pacific Symposium on Biocomputing*, *21*, 504–515.

Orben, A. (2020). Teenagers, screens and social media: A narrative review of reviews and key studies. *Social Psychiatry and Psychiatric Epidemiology*, *55*, 407–414. https://doi.org/10.1007/s00127-019-01825-4

Ouyang, X., Zhou, P., Li, C. H., & Liu, L. (2015, October 26–28). *Sentiment analysis using convolutional neural network* [Conference session]. 2015 IEEE International Conference on Computer and Information Technology, Liverpool, England.

Parry, D. A., Fisher, J., Mieczkowski, H., Sewall, C., & Davidson, B. I. (2022). Social media and well-being: A methodological perspective. *Current Opinion in Psychology*, *45*, Article 101285. https://doi.org/10.1016/j.copsyc.2021.11.005

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., & Duchesnay, É. (2011). Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, *12*(85), 2825–2830. https://www.jmlr.org/papers/v12/pedregosa11a.html

Porturas, T., & Taylor, R. A. (2021). Forty years of emergency medicine research: Uncovering research themes and trends through topic modeling. *The American Journal of Emergency Medicine*, *45*, 213–220.

Ranney, M., Lehrbach, K., Scott, N., Nugent, N., Riese, A., Huang, J., Fong, G., & Rosen, R. (2020, January 7–10). *Insights into adolescent online conflict through qualitative analysis of online messages* [Conference session]. Proceedings of the 53rd Hawaii International Conference on System Sciences, Maui, HI, United States.

Reinecke, L., & Trepte, S. (2014). Authenticity and well-being on social network sites: A two-wave longitudinal study on the effects of online authenticity and the positivity bias in SNS communication. *Computers in Human Behavior*, *30*, 95–102. https://doi.org/10.1016/j.chb.2013.07.030

Schønning, V., Hjetland, G. J., Aarø, L. E., & Skogen, J. C. (2020). Social media use and mental health and well-being among adolescents: A scoping review. *Frontiers in Psychology*, *11*, Article 1949. https://doi.org/10.3389/fpsyg.2020.01949

Schreurs, L., & Vandenbosch, L. (2021). Introducing the Social Media Literacy (SMILE) model with the case of the positivity bias on social media. *Journal of Children and Media*, *13*(3), 320–337. https://doi.org/10.1080/17482798.2020.1809481

Severyn, A., & Moschitti, A. (2015, August 9–13). *Twitter sentiment analysis with deep convolutional neural networks* [Conference session]. Proceedings of the 38th International ACM SIGIR Conference on Research and Development in Information Retrieval, Santiago, Chile.

Song, Y., Pan, S., Liu, S., Zhou, M. X., & Qian, W. (2009, November 2–6). *Topic and keyword re-ranking for LDA-based topic modeling* [Conference session]. Proceedings of the 18th ACM Conference on Information and knowledge management, Hong Kong.

Surkyn, H., Vandekerckhove, R., & Sandra, D. (2021). The impact of analogical effects and social factors on the spelling of partially homophonous verb forms in informal social media writing. *Written Language & Literacy*, *24*(1), 1–37. https://doi.org/10.1075/wll.00046.sur

Tov, W. (2018). Well-being concepts and components. In E. Diener, S. Oishi, & L. Tay (Eds.), *Handbook of well-being*. DEF Publishers. https://doi.org/nobascholar.com

Valkenburg, P. M. (2022). Social media use and well-being: What we know and what we need to know. *Current Opinion in Psychology*, *45*, Article 101294. https://doi.org/10.1016/j.copsyc.2021.12.006

Valkenburg, P. M., Meier, A., & Beyens, I. (2022). Social media use and its impact on adolescent mental health: An umbrella review of the evidence. *Current Opinion in Psychology*, *44*, 58–68. https://doi.org/10.1016/j.copsyc.2021.08.017

Valkenburg, P. M., & Peter, J. (2013). The differential susceptibility to media effects model. *Journal of Communication*, *63*(2), 221–243. https://doi.org/10.1111/jcom.12024

Valkenburg, P. M., Pouwels, J. L., Beyens, I., van Driel, I. I., & Keijsers, L. (2021). Adolescents' experiences on social media and their self-esteem: A person-specific susceptibility perspective. *Technology, Mind, and Behavior*, *2*(2), 1–12. https://doi.org/10.1037/tmb0000037

Vallat, R. (2018). Pingouin: Statistics in Python. *Journal of Open Source Software*, *3*(31), Article 1026. https://doi.org/10.21105/joss.01026

van Atteveldt, W., van der Velden, M. A. C. G., & Boukes, M. (2021). The validity of sentiment analysis: Comparing manual annotation, crowd-coding, dictionary approaches, and machine learning algorithms. *Communication Methods and Measures*, *15*(2), 121–140. https://doi.org/10.1080/19312458.2020.1869198

van Driel, I. I., Giachanou, A., Pouwels, J. L., Boeschoten, L., Beyens, I., & Valkenburg, P. M. (2022). Promises and pitfalls of Instagram data donations. *Communication Methods and Measures*, *16*(4), 266–282. https://doi.org/10.1080/19312458.2022.2109608

van Driel, I. I., Pouwels, J. L., Boeschoten, L., Beyens, I., & Valkenburg, P. M. (2019). *Posting, scrolling, chatting, and Snapping: Youth (14–15) and social media in 2019*. Center for Research on Children, Adolescents, and the Media (CcaM), University of Amsterdam.

Verbeij, T., Pouwels, J. L., Beyens, I., & Valkenburg, P. M. (2022). Experience sampling self-reports of social media use have comparable predictive validity to digital trace measures. *Scientific Reports*, *12*(1), Article 7611. https://doi.org/10.1038/s41598-022-11510-3

Waterloo, S. F., Baumgartner, S. E., Peter, J., & Valkenburg, P. M. (2018). Norms of online expressions of emotion: Comparing Facebook, Twitter, Instagram, and WhatsApp. *New Media & Society*, *20*(5), 1813–1831. https://doi.org/10.1177/1461444817707349

Weinstein, E. (2018). The social media see-saw: Positive and negative influences on adolescents' affective well-being. *New Media & Society*, *20*(10), 3597–3623. https://doi.org/10.1177/1461444818755634

Zeman, J., Cassano, M., Perry-Parrish, C., & Stegall, S. (2006). Emotion regulation in children and adolescents. *Journal of Developmental & Behavioral Pediatrics*, *27*(2), 155–168. https://doi.org/10.1097/00004703-200604000-00014

Author Biographies

Tim Verbeij is a doctoral candidate in the Amsterdam School of Communication Research (ASCoR) at the University of Amsterdam. His research focuses on the validity of social media use measurement methods and the effects of social media use on adolescents' well-being. He is also interested in novel methods to study social media use, for example through social media data donations.

Ine Beyens is an assistant professor in the Amsterdam School of Communication Research (ASCoR) at the University of Amsterdam. Her research focuses on the effects of media on the emotional, psychological, and social development of children and adolescents.

Damian Trilling is a professor of journalism studies at Vrije Universiteit Amsterdam. His research interests include the dissemination of societally relevant information in changing media environments as well as the development and application of computational methods in communication science.

Patti M. Valkenburg is a distinguished professor in the Amsterdam School of Communication Research (ASCoR) at the University of Amsterdam. Her research focuses on adolescents' and adults' individual susceptibility to the affective and cognitive effects of (social) media use.