It's going to be okay:
Measuring Access to Support in Online Communities

Zijian Wang
Symbolic Systems Program
Stanford University
zijwang@stanford.edu

David Jurgens
School of Information
University of Michigan
jurgens@umich.edu

Abstract

People use online platforms to seek out support for their informational and emotional needs. Here, we ask what effect does revealing one's gender have on receiving support. To answer this, we create (i) a new dataset and method for identifying supportive replies and (ii) new methods for inferring gender from text and name. We apply these methods to create a new massive corpus of 102M online interactions with gender-labeled users, each rated by degree of supportiveness. Our analysis shows wide-spread and consistent disparity in support: identifying as a woman is associated with higher rates of support—but also higher rates of disparagement.

1 Introduction

Despite substantial efforts to reduce gender disparities in online social contexts, gender gaps persist and, increasingly, negatively affect women through online harassment (Duggan, 2017). Online social platforms still serve a critical role for individuals as they seek to fill informational and emotional needs, frequently by interacting with others (Goswami et al., 2010; Chuang and Yang, 2012; Hether et al., 2016). The supportive replies of others help promote personal well-being (MacGeorge et al., 2011), yet unsupportive replies can not only lead to distress but discourage online engagement altogether. Given gender disparity in the receipt of anti-social behavior, to what degree does this disparity persist in individuals' receipt of support? We answer this question, illustrated in Figure 1, by examining supportive and unsupportive message rates across millions of online interactions, using a new computational model of support. Our work is motivated by an agenda of promoting supportive online platforms where people can participate equally.

This work connects with the growing body of

> **Comment**: `KatieZ22`: I'm nervous about my differential calculus exam next week. My current idea is work through problems on previous exams. But yiiiikes.
> ∟ **Reply**: `PizzaMagic`: You can ace that test! Your plan seems smart and you have plenty of time to prepare. :)

Figure 1: In this fictitious example, `KatieZ22` receives a supportive reply from `PizzaMagic`. In choosing their names, each user has chosen a particular gender performance, signaling female and gender-anonymity, respectively. In online settings, such gender performances evoke stereotypes that affect how others interact and provide access to online resources. Our study asks what effect does this gender signaling have on individuals receiving support and disparagement?

computational studies of gender disparity in online behavior (e.g., Lam et al., 2011; Magno and Weber, 2014; Garimella and Mihalcea, 2016; Li et al., 2018); our work here examines this disparity along a new dimension, support, and unlike prior work, examines disparity along the full spectrum of both pro-social (supportive) and anti-social (unsupportive) behaviors. Prior works have also examined the language of support in online support forums for health-related issues (Biyani et al., 2014; Wang et al., 2012; De Choudhury and De, 2014; Althoff et al., 2016; De Choudhury and Kiciman, 2017), often with the aim of improving people's access. Here, we aim to study support in general, everyday interactions, drawing upon theories of how support is expressed in language (Cutrona and Suhr, 1992; Wright et al., 2003).

Our investigation provides four main contributions. First, we introduce a new task of rating the supportiveness of a message and provide an accompanying dataset of 9,032 post-reply pairs with annotations (§2). Second, using this data, we develop a new computational model for automatically identifying supportive and unsupportive replies (§3), using theory-based features that operationalize linguistic strategies for giving support. Third, we develop a new state-of-the-art sys-

tem for classifying the gender of a username (§4) and construct a massive dataset of over 102M post-reply pairs from three online platforms (§5), where participants are labeled by gender. Further, the text of each post is rated for its gender predictiveness, enabling studying gender performance at the name and textual levels. Finally, we apply our support classifier to our social interaction dataset to reveal wide-spread disparity on the basis of gender (§6). Our results show that when gender is performed, female performances are associated *both* with higher rates of supportive comments *and* with higher rates of unsupportive comments, highlighting that gender disparity is not just for negative behaviors online.

2 The Language of Support

Individuals engage in online platforms for a variety of reasons and supportive responses to this engagement can take many forms (Shumaker and Brownell, 1984; Vaux, 1985), from informational support like advice to emotional support like expressions of sympathy. Responders may choose from different linguistic support strategies depending on the speaker and context (Cutrona and Suhr, 1992). For example, given an individual commenting on a Wikipedia talk page about their idea for adding new content, a responder may point to an additional resource they can use, whereas given an individual posting to Reddit for relationship advice, a responder may express sympathy. Our goal is to study the language and behavior of *everyday* supportive or unsupportive interactions as they occur on three large social platforms: Reddit, StackExchange, and Wikipedia. These platforms represent common settings people seek out to engage in discussions and ask for help. Therefore, we annotate a dataset by degrees of support and analyze how linguistic expectations of support manifest in online interactions.

Data and Annotation Post and reply pairs were selected from the three platforms. Many of these interactions are short and therefore to increase diversity, annotated pairs were sampled by balancing by platform and the lengths of posts and replies seen in each.

As a social activity, support is often expressed by drawing upon other social strategies such as politeness (Feng et al., 2013). To help focus annotators' attention on support specifically and disentangle related social cues, we pair our sup-
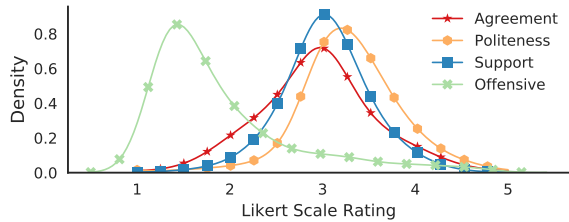

Figure 2: Annotators' rating distributions.

Rating	Example Reply
1.33	see your arse mate, stop talking out of it.
2.0	We dont like your kind.
3.0	I was referring to <link>
3.33	thanks chief.
3.66	Not many people, apparently, but I'm listening!
4.0	Im in as well! Meet up at the Tavern?!
4.33	I like your style!
4.66	Love this. Thank you, more please.

Table 1: Examples of annotator ratings of Support from 1 to 5. Additional examples are in Supplemental §2.

port annotation with contrastive annotation questions for three other related phenomena: agreement (with the post's message), politeness, and offensiveness. Annotators were asked to rate support on a five-point Likert scale from very unsupportive and very supportive, with analogous questions each for agreement and politeness; offensiveness was rated on a five point scale from inoffensive to very offensive.

All data was annotated using CrowdFlower with detailed instructions and example replies for each level of the Likert scale. Each task presented five post-reply pairs and with detailed instructions that ask annotators to focus on rating each reply along these four dimensions. Annotators were required to pass a training phase where they had at least 70% agreement with a gold standard annotation on 10 items. After training, each task included one control question, which was used to remove annotators whose agreement with the gold standard fell below 70%.

In total, 9,032 instances were annotated by three annotators, who had a Krippendorff's α of 0.766, indicating substantial agreement on the data (Artstein and Poesio, 2008). Figure 2 shows the distribution of ratings. Support is positively correlated with agreement (r=0.71) and politeness (r=0.51), though annotators rated replies as having more politeness than support on average. Support annotation examples are shown in Table 1. Offensiveness was negatively correlated (r=-0.38) with support.

Analysis Supportive replies can use a variety of strategies for offering support, depending on need

	Strategy	Support	In Top-25%	Example
INFORMATIONAL	Suggestion advice	0.043	16.3%	You might try...
	Referral	0.091	1.3%	Please see [URL]
	Situational appraisal	−0.071**	1.5%	Your situation sounds like...
	Teaching	−0.065***	2.7%	The reason that's happening...
TANGIBLE	Direct offer to do something	−0.020	0.04%	Do you want me to?
	Willingness	0.266***	1.6%	I could help you...
ESTEEM	Compliment	0.337***	23.8%	Great idea!
	Validation	0.248***	27.2%	You're right about...
	Relief of blame	0.490***	1.2%	It's not your fault that...
	Companionship Reminder	−0.089*	0.7%	Your friends and family still...
EMOTIONAL	Sympathy	0.041	0.1%	Sorry to hear that
	Listening	−0.104***	3.1%	Why did you feel...
	Empathy	0.067	0.08%	I know how you feel...
	Encouragement	0.423***	1.6%	Go for it
	Accommodation	0.035	74.7%	
	Emotion	0.081***	27.2%	

Table 2: Support strategies and their presence in our data, as shown through the mean supportiveness rating (-2 to 2) for replies using that strategy and the percentage of posts in the top 25% of the most-supportive that employ the strategy. For supportiveness, posts are compared with all others not employing that strategy, with significance measured using the Mann-Whitney U test. Throughout the paper, *** denotes p<0.001, ** p<0.01, and * p<0.05.

and context. Cutrona and Suhr (1992) proposed a broad taxonomy of support strategies based on in-person interactions, such as offering an appraisal of the current situation or seeking to relieve the other person of blame. We examine to what degree are these strategies employed in online, relatively-anonymous settings and whether their usage online is associated with higher perceived support. To test this, support strategies were automatically identified using a combination of regular expressions for lexical patterns and dependency-parsed trees, together with specialized lexicons matching each strategy and rules for detecting negation. For example, suggestions were detected by identifying a second-person subject with a modal verb indicating possibility (Quirk et al., 1985, p. 219).

Many of the strategies suggested by Cutrona and Suhr (1992) for expressing support in person were also observed online; Table 2 shows the average supportiveness rating for replies containing each strategy, where supportiveness is centered to $[−2, 2]$. Further, their effect on perception of supportiveness, while small, is significant and positive for many. However, we do observe two notable negative trends. First informational support strategies of aiding a person by reassessing their situation (e.g., offering a new perspective) and by teaching were considered less supportive. We observed that in several cases these strategies

were employed when an individual was not seeking support, in which case unrequested new information can appear condescending. Indeed, replies employing these two support strategies were still rated polite, 0.30 and 0.31 mean politeness respectively, suggesting the context in which new information is given weighs heavily on whether it is treated as supporting the individual.

Support can also be conveyed implicitly through unconscious stylistic choices. Rains (2016) notes linguistic accommodation is frequently observed in supportive responses, where individuals match the function word frequency of the original communication (Bucholtz and Hall, 2005; Danescu-Niculescu-Mizil et al., 2011). Shown in the bottom of Table 2, we observe only a weak non-significant positive association between support and accommodation (as measured using Danescu-Niculescu-Mizil et al. (2011)), though the single post-reply unit of analysis limits our ability to detect long-term accommodation across multiple dialog turns. Liviatan et al. (2008) and Li and Feng (2014) found that supportive replies often to contain more emotional language, which evokes a personal connection and intensity (Spottswood et al., 2013; Braithwaite et al., 1999). Shown in the bottom of Table 2, we find a significant and positive association where more emotive posts are viewed as more supportive.

Finally, we note that a few strategies suggested by Cutrona and Suhr (1992) were not seen in our annotated dataset. Strategies of offering to participate, using physical affection, assurances of confidentiality, material and financial loans, and prayer were not seen; a broader scan of our unannotated data did find these attested but rare in practice. We attribute this rareness to the public, online nature of the interactions, in contrast to the interpersonal setting studied by Cutrona and Suhr (1992).

3 Computational Model of Support

Our primary objective is to measure the relationship between identity and support in online social systems. Therefore, we next develop a classifier to automatically label replies with support.

Features As a social activity, the language of support draws upon multiple lexical and stylistic cues. We base on classifier on *theory-inspired* and data-driven features. The first set consists of the operationalized linguistic strategies for expressing support, shown in Table 2. Further, in constructing our feature set, we build upon past linguistic analyses of related social-situated language. Wellman and Wortley (1990) note that the availability of support is related to social distance, which is in part expressed linguistically through the degree of formality (Hovy, 1987; Sigley, 1997). Therefore, we include features from Pavlick and Tetreault (2016), which examined linguistic markers of formality. Advice giving is a core component of many theories of support (MacGeorge et al., 2011) and such advice is frequently wrapped in politeness language (Feng et al., 2013), e.g., hedging suggestions rather than imposing direction, which provides face-saving opportunities for the person receiving support (Clark and Schunk, 1980). Therefore, we include the feature set of Danescu-Niculescu-Mizil et al. (2013), which though focused on requests, provides many general lexical patterns for politeness.

Beyond these, we include features motivated by observational studies of support. In analyzing online support groups, Alpers et al. (2005) found that LIWC (Pennebaker and Stone, 2003) was valid as a construct for analyzing messages and compared similarly human judgments about the categories. To capture emotional language (Li and Feng, 2014; Liviatan et al., 2008), we include the NRC emotion lexicons of Mohammad and Turney (2013). Given the informational support strategy,

we include lexicons from argumentation for capturing explanatory replies (Teufel, 2000).

Support may be given in response to stressors, which change in nature throughout a person's lifetime (Vaux, 1985; Segrin, 2003). To potentially capture variation in the language of support based on the posting individuals, we include features known to be associated with age such as elongation and capitalization (Goswami et al., 2009; Barbieri, 2008), grammatical differences in sentence construction and length (Hovy and Søgaard, 2015), and a lexicon for age of acquisition (Kuperman et al., 2012).

Data-driven features include (1) lexical features capturing the presence of n-grams, their relative frequency, (2) grammatical features from dependency-parsed triples, which are also backed off to parts of speech, (3) word lexicons for formality, sentiment, and subjectivity, (4) style features such as word and sentence length, complexity, and use of contractions, and (5) the average word vector for the sentence.

In total, our model includes 23,903 features, the bulk of which are n-grams and dependency triples. A detailed listing of all features is provided in Supplemental §1.

One notable feature that we did not include was the presence of self-disclosure in a reply, which has been linked to high-social support as a way of conveying connection and empathy (Wright et al., 2003). While computational models for self-disclosure have been proposed (Bak et al., 2014; De Choudhury and De, 2014), we were unable to scale these methods to the size of our analysis.

Task Setup Support ratings are discretized to create a ternary classification task with labels $\{-1, 0, 1\}$, denoting unsupportive, neutral, and supportive comments. Ratings were discretized by treating all those ratings $\leq$-0.67 as negative and those with a rating $\geq$0.67 as positive.

A Random Forest classifier was trained on all 23,903 features; random forests are robust to overfitting even with large numbers of features, making them suitable for this high-dimensional feature space (Fernández-Delgado et al., 2014). Furthermore, random forests are able to learn conjunctive features, allowing us to learn how combinations of strategies are employed to yield support. As the majority of posts are neutral, we mitigate the class imbalance using SMOTE (Chawla et al., 2002) to generate synthetic examples in the

Model	Macro-F1
Our Model	**0.52**
Bigram Features	0.43
Unigram Features	0.42
Support Features (Table 2)	0.40
Majority	0.29
Random	0.26

Table 3: Support classification performance

training fold using the 5 nearest neighbors, taking care to avoid contamination of the test set. The classifier is implemented using Scikit-learn (Pedregosa et al., 2011) and syntactic processing was done using spaCy (Honnibal and Johnson, 2015). Word vectors are the publicly released Google-News word2vec vectors (Mikolov et al., 2013).

Three works have examined related tasks where Biyani et al. (2014) and Khanpour et al. (2018) classify posts in online cancer support groups as providing informational or emotional support and Wang et al. (2012) classify the degree of support along these dimensions. Here, we solve a more general task that includes unsupportive comments and is in the general domain.

Evaluation We compare our full model for predicting support against three models: our 14 features for detecting support strategies from Table 2, a model trained on the subset of unigram features (4,352), and a model trained on bigram features (8,897), the latter of which is known to be a strong lexical baseline (Wang and Manning, 2012). All models were tested using five-fold cross-validation with Macro-F1 for evaluation and including baselines for labeling instances at random or choosing the most frequent.

Our full model obtains substantial improvements over all baselines and models, as shown in Table 3. Further, the simple support strategy features provide a large and statically-significant improvement over the two baselines. The model using support strategy features performs similarity to the unigram model, despite having two orders of magnitude fewer features. A follow-up analysis on cross-platform performance, described in Supplemental §1.2, showed that while within-platform performance was relatively high (0.54 Macro F1 for Reddit and 0.53 for Wikipedia), performance for the more technical StackExchange site was lower both within (0.44) and across (0.40 when trained on Reddit and 0.42 when on Wikipedia).

Examining our full model's most important features showed that the two support strategies for validation and compliments (cf. Table 2) were

the most important features, followed closely by lexicons for emotion: Anger in LIWC, Disgust in NRC, and the positive sentiment in Liu et al. (2005), all of which were motivated by theory. These results confirm that our theory-inspired features are both salient for supportiveness and effective as features.

4 Inferring Gender

As a part of interacting, individuals present a view of themselves as an interlocutor, revealing aspects about themselves such as gender through explicit means (Marwick, 2013; Allen and Wiles, 2016), e.g., profile pictures, or through implicit—and potentially unconscious—cues such as stylistic choices in language (e.g., Eckert and McConnell-Ginet, 2003; Bamman et al., 2014). In the relatively anonymous and deindividuated online setting, these identity cues can have a profound impact on how other perceive and interact with them (e.g., Mickelson et al., 1995; Herring, 2003; Ammari et al., 2014; Megarry, 2014) and these minimal gender cues give rise to full-fledged social stereotypes and, potentially, the negative behavior that comes when treating someone as a stereotype (Kiesler et al., 1984; Lea and Spears, 1991; Postmes et al., 1998; Wang et al., 2009). Here, we develop methods for inferring gender from two signals: (1) names that users chose; and (2) implicit cues conveyed by linguistic features.

4.1 Gender from Names

Prior work has developed models for inferring gender from username alone (e.g., Tang et al., 2011; Liu and Ruths, 2013; Jaech and Ostendorf, 2015; Knowles et al., 2016). Here, we develop a new character-based neural model that incorporates rich gender-labeled username information for identifying additional gender-salience cues in usernames from roles and attributes, e.g., `SuperDad1` or `AspiringActress99`.

Data Individuals convey their gender in multiple ways beyond using gender-associated names. Therefore, to capture this variety, we collect usernames from two online platforms where users have self-declared their perceived gender. First, Twitter usernames and screen names were collected from a 10% sample from 2014 to 2017. Here, we identify usernames whose biography contains an explicit mention of their gender, e.g., by stating a gendered role "mom to two kids" or spec-

ifying pronoun preferences "he/him/his." Gendered profiles were collected for 4,900,250 individuals using a selection of lexical patterns with aggressive filtering to remove false positives. Second, we collect 283,427 usernames from Reddit identified through self-declarations of gender in /R/RELATIONSHIPS, e.g., "I [23F] need to talk to my boyfriend [27M]", and 84,068 usernames where the user has chosen a gender-indicating flair (a visual icon displayed within the subreddit). These two sources provide much-needed variation for gender in usernames beyond those mirroring full names.

Model Given a username, we infer gender using a character-based encoder consisting of three stacked LSTM networks (Hochreiter and Schmidhuber, 1997). Following platform restrictions on usernames, character sequences are restricted to being in ascii range and are embedded into 16 dimensional vectors as input. Adopting best practices (Ioffe and Szegedy, 2015), batch normalization is applied prior to the dense layer used to compute the gender prediction. LSTMs were sized at 256 after limited hyperparameter tuning on development data. We optimize with Adam (Kingma and Ba, 2014) with a learning rate of 0.002.

Training and Evaluation All data is partitioned into 80% train, 10% development, and 10% test splits. As some usernames are repeated in different communities, we keep only one unique instance prior to partitioning to avoid leakage between partitions.

Training mini-batches were balanced for both genders, which yielded better performance in tests on the development data. We compare the performance of our model on the test set against two current state-of-art systems available off the shelf for inferring gender from usernames, demographer (Knowles et al., 2016) and Jaech and Ostendorf (2015). Demographer is trained on names from the Social Security Administration and the method of Jaech and Ostendorf (2015) is trained on usernames from OkCupid and uses 3.5M Snapchat usernames for self-learning to improve accuracy.

As shown in Table 4, our model outperforms both systems by substantial margins for both Twitter and Reddit data. In tests on data from both papers reported in Supplemental §3, our model also outperforms their systems. High accuracy is not expected for these models in most domains, as many usernames do not signal gender.

Method	Twitter	Reddit
Our Model	**0.7785**	**0.6299**
Jaech and Ostendorf (2015)	0.7028	0.5935
Knowles et al. (2016)	0.6520	0.5216

Table 4: Gender inference (Macro-F1)

4.2 Gender from Text

Gender can also manifest through more subtle, stylistic cues (e.g., Schnoebelen, 2012; Flekova and Gurevych, 2013; Bamman et al., 2014; Volkova et al., 2015; Garimella and Mihalcea, 2016; Carpenter et al., 2016). Thus, even when a person chooses a neutral username, their linguistic style may reveal their gender. Therefore, we construct a regression model to infer the degree to which either gender is expressed through text.

Data and Model Gender-labeled post data was constructed using held-out data from our three platforms where posts were authored by a user with a high-confidence gender prediction. Posts were randomly sampled across forums (e.g., subreddits) from the held-out data to achieve gender parity with 555K posts for Wikipedia, and 58K for StackExchange; Reddit was subsampled to 1M posts total due to its size.

Features were selected by drawing upon prior work: (i) stylistic features like punctuation and number frequencies, casing, word length and (ii) content features including n-grams, sentiment, and specialized lexicons like LIWC. A full listing of features is reported in Supplemental §1.

Following prior work (Bamman et al., 2014), a logistic regression model was trained for each platform using L2 regression; we adopt separate models for each to better adapt to any platform-specific gender variation.

Evaluation Models were evaluated using AUC with five-fold cross validation, with 0.661 AUC for StackExchange, 0.700 for Wikipedia, and 0.661 for Reddit. The models perform substantially better than random choice (0.5) for the challenging task of inferring gender from a single post, as many posts contain no signal of gender. Additional analyses are reported in Supplemental §1.2.

5 Gender-Salient Interaction Data

To quantify the social support people receive online, we examine communications from three major online communities: Reddit, StackExchange, and Wikipedia. We refer to a communication between individuals as a *post* with a *reply*, defining

	Male	Female	Neutral
Reddit	1,017,455	213,849	1,211,813
SE	2,431,234	289,864	836,531
Wiki.	80,876	14,073	46,954

Table 5: Users with high-confidence or neutral gender

Gender: Female	Gender: Male	Gender: Neutral
Shanakitty	AdamMcAdamson	kazkeb
cassycas	ChipCarlson11	xuebin671
kelseylenae	BaBa_Dad	ConfigurationalYes
Mrs_BruceWayne	BarryCA67	thelizardof_Oz
madelaine00x	BenJewish	oibird
norma-gaspard	dojoguy	Merpageddon

Table 6: High-confidence and neutral name examples.

these for each platform next.

Reddit Reddit data was selected using a longitudinal sample of one month (July) per year, from 2006 to 2017, and a continuous sample of one full year's data in 2017. Our initial data consists of the top 10,000 subreddits ordered by the total number of posts in the data. From these communities we restrict our analysis to a comment and its first reply, which reduces confounds from multi-party communication. These post-reply pairs were further filtered to remove non-English posts using Google CLD2 (McCandless, 2010), yielding 434.29M candidate communications.

StackExchange StackExchange (SE) contains substantial social interaction in the comment to posts and replies (Ahn et al., 2013; Danescu-Niculescu-Mizil et al., 2013). These communications often expand beyond the immediate topic. Directed communication within these comments is frequently signaled using an explicit mention starting with an "@," which we use to identify pairs. In total, we collected post-reply pairs from the full history of all StackExchange, yielding 3.16M pairs across 162 sites.

Wikipedia Wikipedia features an active social component in its talk pages, with more personal communication–or even personal attacks–during debates around appropriateness or suggested changes (Bender et al., 2011). Similar to Reddit, we construct post-reply pairs by identifying each comment and its first response on a talk page, yielding 26.7M pairs from 387K talk pages.

Assigning Gender All post-reply pairs were labeled using our post classifier (§4.1). Posts with *high-confidence* gender predictions (softmax probability > 0.9 or < 0.1) were labeled with the predicted gender. To contrast the effects of having

a gendered name, we construct a complementary dataset where the posting user's username is effectively gender *neutral*, e.g., `user1209`; these neutral names are chosen from those with near-chance probability in the output softmax $0.45 < p < 0.55$. The relative counts of high-confidence and neutral gender names in each platform are shown in Table 5, along with examples in Table 6.

Restricting the dataset to pairs where we have a salient identity, our final dataset for analysis consists of 49.58M, 0.72M, and 3.69M pairs for gender in Reddit, StackExchange, and Wikipedia; and 46.19M, 201.7K, and 1.60M for neutral in each, respectively. Where possible, we also record any high-salience identities for replying users.

6 Gender and Support

The gender cues provided through computer mediated communication provide enough information that a person will fill in the result with a stereotype (Lea and Spears, 1991; Spears and Lea, 1992). What effect might this stereotyping have for access to support? While establishing full causality for an answer is infeasible in our current observational study, we take the first step by quantifying whether disparity in support exists and examine what contextual factors may affect support giving. Using our classifier, we label the 102M post-reply pairs from our dataset (§5), which includes both high-confidence and gender-neutral users.

Model To quantify access to support, we construct separate mixed-effect logistic regression models for predicting the dependent variable of whether a post will receive a supportive reply and for whether it will receive an unsupportive reply. Random effects are added for each community within a platform, which capture the variance in support rates between communities, e.g., due to differences in community norms, topic, or size. As fixed effects, we include a categorical variable for name-inferred gender, always using the gender-neutral condition as the reference coding, which, critically, allows us to examine the changes of support for revealing gender relative to users whose identity is effectively anonymous. We include the predicted probability of a user's gender from their writing (§4.2), centered to $[-0.5, 0.5]$ such that 0 denotes a gender-neutral writing style. Finally, we include interaction terms for writing and names to capture effects of joint gender performance.

Results Three main results are observed. First,

	SUPPORTIVE			UNSUPPORTIVE		
	Reddit	**StackExchange**	**Wikipedia**	**Reddit**	**StackExchange**	**Wikipedia**
intercept	−2.3775***	−2.4729***	−2.9298***	−3.0232***	−4.3734***	−2.9787***
♀name	0.5480***	0.5409***	0.7376***	0.2927***	0.6071***	0.3543***
♂name	0.4652***	0.6639***	0.6422***	0.3230***	0.6445***	0.3313***
♀✎	1.1174***	−1.0778***	0.5261***	0.2520*	0.7398***	−0.1451**
♀name ∧ ♀✎	0.8802***	1.2751***	0.1155	0.2662*	0.0537	0.1695
♂name ∧ ♀✎	0.0535	0.8119***	−0.2070***	0.2016	0.3556*	0.4750***

Table 7: Logistic regression coefficients for predicting whether a post will receive a supportive reply (left) or unsupportive reply (right) on the basis of a gendered name and writing (denoted ✎), with coefficients for name and writing interactions. Name gender is categorical with the reference coding is neutral. The writing coefficient is center at 0 (neutral) with positive values being more female.

the use of a gender-conveying name is associated with both higher rates of supportive comments and unsupportive comments. Our results agree with those from the small scale study of Feng et al. (2013) who found that accounts with human pictures and person-sounding usernames receive higher social support. Indeed, while several studies have touted the benefits of anonymity online for discussing sensitive topics (Campbell and Wright, 2002; Wright, 2002a,b), our results suggest that selecting a gender neutral name may lead to lower support overall. Our findings also reinforce the observation that the personal-anonymity online does not lead to equal support due to cues about identity (Postmes and Spears, 2002).

Second, the rates of supportive and unsupportive comments are significantly associated with *both* kinds of gender performances (i.e., names and writing style). These results suggest that online audiences are sensitive to both kinds of overt and implicit gender displays and that even innocuous choices such as gendered usernames can shape our online interactions.

Third, when gender is performed in together name and writing, female performances are consistently associated across all three platforms with higher rates of receiving supportive replies and unsupportive replies. Note that this trend is seen in the cumulative effect on support after combining coefficients for the interactions term with the coefficients for writing and name. We illustrate this cumulative effect for two types of gender performances in Reddit, shown as separate axes in Figure 3. Indeed, when a user has a gendered username, the cumulative effect of male writing performance is consistently associated with fewer supportive replies. In small-scale interpersonal studies, Abbey et al. (1991) and Barbee et al. (1993) note men are more likely to receive unsup-
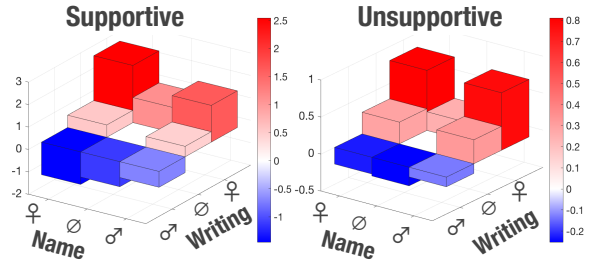

Figure 3: Cumulative effect of name and writing performances on the log-odds coefficients for receiving comments of each type in Reddit when the gender is maximally performed (cf. Table 7), excluding the intercept effect for clarity (∅ denotes neutral performance). Positive values (red) denote increased comment frequency and negative values (blue) show decreased frequency. This plot shows that as female performance becomes more salient (shown bottom to top visually), such users receive increasingly-higher rates of supportive and unsupportive replies in Reddit.

portive comments; however, we did not observe this disparity in our online setting.

Does the replier's gender matter? Mickelson et al. (1995) note that men and women differ in *how* they receive support, with the gender composition of the interacting pair driving the kind of supportive behavior. Here, we examine whether men and women differ in the *rates* they give support to one another, using gendered names as a proxy for identity. Because only Reddit has sufficient data, we construct a mixed-effect regression model for Reddit using the 4.5M post-reply pairs where the replier has a high-confidence or neutral gender and include the replier's gender as a factor with interactions for the poster. The results shown in Table 8 reveal two main conclusions. First, men and women give supportive comments at different rates, with women being far more likely to leave supportive replies and less likely to leave unsupportive replies. Second, the interaction terms show that there is minimal dyadic interaction between the gender identities of the poster and replier with

	Sup.	Unsup.
intercept	−2.391***	−3.107***
P:♀name	0.561***	0.288***
P:♂name	0.450***	0.330***
P:♀✎	1.263***	0.290***
R:♀name	0.249***	−0.153***
R:♂name	−0.057***	−0.036**
P:♀name ∧ P:♀✎	0.914***	0.212**
P:♂name ∧ P:♀✎	0.082	0.131
P:♀name ∧ R:♀name	−0.103***	−0.037
P:♂name ∧ R:♀name	−0.033	0.026
P:♀name ∧ R:♂name	−0.004	0.036*
P:♂name ∧ R:♂name	0.031*	−0.003
P:♀✎ ∧ R:♀name	0.209	−0.202
P:♀✎ ∧ R:♂name	−0.331**	−0.228*
P:♀name ∧ P:♀✎ ∧ R:♀name	−0.343	0.376
P:♀name ∧ P:♀✎ ∧ R:♂name	0.329	0.312
P:♂name ∧ P:♀✎ ∧ R:♀name	0.067	0.178
P:♂name ∧ P:♀✎ ∧ R:♂name	0.024	0.203

Table 8: Regression coefficients for Reddit when the gender identity of the replier is known. The post author's identity is denoted with a P and replier's with R.

respect to rates of giving supportive or unsupportive comments. We only observe significant interactions indicating (1) replying users with female names are *less* likely to leave supportive replies to posting users with female names and (2) replying users with male names are *i*) *more* likely to leave supportive replies to other users with male names, *ii*) *less* likely to leave supportive comments if the writing appears more female, *iii*) *more* likely to leave unsupportive comments if the posters name is female, and *iv*) *less* likely to leave unsupportive comments if the writing appears more female.

Limitations The observations of our study should still be viewed within its practical limitations, of which we note two. First, in examining the content of replies, we do not control for potential direct or indirect requests for help in text that may ultimately affect the rates of support. This issue could be a potential confound, as the cultural norms for masculinity often promote self reliance (Addis and Mahalik, 2003), ultimately leading to gendered differences in requests. Second, this observational study cannot establish causality between gender displays and support; while the disparity is real, exogenous factors could potentially explain the disparity without finding gender displays as a cause, though the mixed effects still control for some contextual variability in the different support frequencies across communities. In spite of these limitations, we view this work as an important first step for demonstrating gender disparity in support—both positive and negative—and inviting future work to establish a causal explanation.

7 Ethical Considerations

The use of gender as a variable in NLP requires that we also discuss ethical considerations resulting from this work, as it directly relates to identity and the dignity of persons being studied. Following the guidelines of Larson (2017) for using gender in NLP, our use of gender is intentional and central to this study on gender disparities in received support. We base our notion of gender as one of linguistic performance (DeFrancisco et al., 2013), in which individuals adapt their style and name to emphasize or de-emphasize certain aspects of their gender identity (Eckert, 2008). Accordingly, we have opted represent gender performance along a graded scale, though we recognize that this representation does not capture non-binary gender identities.

The gender inference methods introduced here raise ethical considerations as they ultimately enable automatic identification of gender for any person on the basis of name or writing (Hamidi et al., 2018). Such technology could be used to unfairly identify and target persons of either gender for malicious behavior or may harm through misgendering individuals. Ultimately, we decided that such risk was acceptable given the positive impact of our study on revealing gender disparity. We hope to also use our method to better support privacy-preserving behavior (Allen and Wiles, 2016; Reddy and Knight, 2016) by helping individuals identify and change names or statements that would indicate a particular gender. Further, we hope that when used in combination with our support classifier and a larger context of gendered interactions (Voigt et al., 2018), these technologies can identify healthy communities that are supportive of all people.

8 Conclusion

Individuals use social media to support their informational and emotional needs. Our study has shown wide-spread disparity in the levels of support individuals receive on the basis of their perceived gender. Our results were made possible through the development of a new massive 102M post-reply dataset tagged with high-salience and neutral gender and the introduction of a new task, annotated dataset, and model for classifying supportive messages. All data, code, and annotation guidelines are publicly released at `https://github.com/davidjurgens/support`.

Acknowledgments

References

Antonia Abbey, Frank M Andrews, and L Jill Halman. 1991. The importance of social relationships for infertile couples' well-being. In *Infertility*, pages 61–86. Springer.

Michael E Addis and James R Mahalik. 2003. Men, masculinity, and the contexts of help seeking. *American Psychologist*, 58(1):5.

June Ahn, Brian S Butler, Cindy Weng, and Sarah Webster. 2013. Learning to be a better q'er in social q&a sites: Social norms and information artifacts. *Proceedings of the Association for Information Science and Technology*, 50(1):1–10.

Ruth ES Allen and Janine L Wiles. 2016. A rose by any other name: Participants choosing research pseudonyms. *Qualitative Research in Psychology*, 13(2):149–165.

Georg W Alpers, Andrew J Winzelberg, Catherine Classen, Heidi Roberts, Parvati Dev, Cheryl Koopman, and C Barr Taylor. 2005. Evaluation of computerized text analysis in an internet breast cancer support group. *Computers in Human Behavior*, 21(2):361–376.

Tim Althoff, Kevin Clark, and Jure Leskovec. 2016. Large-scale analysis of counseling conversations: An application of natural language processing to mental health. *Transactions of the Association for Computational Linguistics (TACL)*, 4:463.

Tawfiq Ammari, Sarita Yardi Schoenebeck, and Meredith Ringel Morris. 2014. Accessing social support and overcoming judgment on social media among parents of children with special needs. In *Proceedings of the International Conference on Web and Social Media (ICWSM)*.

Ron Artstein and Massimo Poesio. 2008. Inter-coder agreement for computational linguistics. *Computational Linguistics*, 34(4):555–596.

JinYeong Bak, Chin-Yew Lin, and Alice Oh. 2014. Self-disclosure topic model for classifying and analyzing twitter conversations. In *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 1986–1996.

David Bamman, Jacob Eisenstein, and Tyler Schnoebelen. 2014. Gender identity and lexical variation in social media. *Journal of Sociolinguistics*, 18(2):135–160.

Anita P Barbee, Michael R Cunningham, Barbara A Winstead, Valerian J Derlega, Mary R Gulley, Pamela A Yankeelov, and Perri B Druen. 1993. Effects of gender role expectations on the social support process. *Journal of Social Issues*, 49(3):175–190.

Federica Barbieri. 2008. Patterns of age-based linguistic variation in American English. *Journal of Sociolinguistics*, 12(1):58–88.

Emily M Bender, Jonathan T Morgan, Meghan Oxley, Mark Zachry, Brian Hutchinson, Alex Marin, Bin Zhang, and Mari Ostendorf. 2011. Annotating social acts: Authority claims and alignment moves in wikipedia talk pages. In *Proceedings of the Workshop on Languages in Social Media*, pages 48–57. Association for Computational Linguistics.

Prakhar Biyani, Cornelia Caragea, Prasenjit Mitra, and John Yen. 2014. Identifying emotional and informational support in online health communities. In *Proceedings of the International Conference on Computational Linguistics (COLING)*, pages 827–836.

Dawn O Braithwaite, Vincent R Waldron, and Jerry Finn. 1999. Communication of social support in computer-mediated groups for people with disabilities. *Health Communication*, 11(2):123–151.

Mary Bucholtz and Kira Hall. 2005. Identity and interaction: A sociocultural linguistic approach. *Discourse Studies*, 7(4-5):585–614.

Kristen Campbell and Kevin B Wright. 2002. On-line support groups: An investigation of relationships among source credibility, dimensions of relational communication, and perceptions of emotional support. *Communication Research Reports*, 19(2):183–193.

Jordan Carpenter, Daniel Preotiuc-Pietro, Lucie Flekova, Salvatore Giorgi, Courtney Hagan, Margaret L Kern, Anneke EK Buffone, Lyle Ungar, and Martin EP Seligman. 2016. Real men don't say "cute": Using automatic language analysis to isolate inaccurate aspects of stereotypes. *Social Psychological and Personality Science*.

Nitesh V Chawla, Kevin W Bowyer, Lawrence O Hall, and W Philip Kegelmeyer. 2002. SMOTE: synthetic minority over-sampling technique. *Journal of Artificial Intelligence Research (JAIR)*, 16:321–357.

Katherine Y Chuang and Christopher C Yang. 2012. Interaction patterns of nurturant support exchanged in online health social networking. *Journal of Medical Internet Research*, 14(3).

Herbert H Clark and Dale H Schunk. 1980. Polite responses to polite requests. *Cognition*, 8(2):111–143.

Carolyn E Cutrona and Julie A Suhr. 1992. Controllability of stressful events and satisfaction with spouse support behaviors. *Communication Research*, 19(2):154–174.

Cristian Danescu-Niculescu-Mizil, Michael Gamon, and Susan Dumais. 2011. Mark my words!: linguistic style accommodation in social media. In *Proceedings of the International Conference on World Wide Web (WWW)*, pages 745–754. ACM.

Cristian Danescu-Niculescu-Mizil, Moritz Sudhof, Dan Jurafsky, Jure Leskovec, and Christopher Potts. 2013. A computational approach to politeness with application to social factors. In *Proceedings of the Annual Meeting of the Association for Computational Linguistics (ACL)*.

Munmun De Choudhury and Sushovan De. 2014. Mental health discourse on reddit: Self-disclosure, social support, and anonymity. In *Proceedings of the International Conference on Web and Social Media (ICWSM)*.

Munmun De Choudhury and Emre Kiciman. 2017. The language of social support in social media and its effect on suicidal ideation risk. In *Proceedings of the International Conference on Web and Social Media (ICWSM)*, pages 32–41.

Victoria Pruin DeFrancisco, Catherine Helen Palczewski, and Danielle D McGeough. 2013. *Gender in communication: A critical introduction*. Sage Publications.

Maeve Duggan. 2017. Online harassment 2017. *Pew Research Center*.

Penelope Eckert. 2008. Variation and the indexical field. *Journal of Sociolinguistics*, 12(4):453–476.

Penelope Eckert and Sally McConnell-Ginet. 2003. *Language and gender*. Cambridge University Press.

Bo Feng, Siyue Li, and Na Li. 2013. Is a profile worth a thousand words? how online support-seeker's profile features may influence the quality of received support messages. *Communication Research*, 43(2):253–276.

Manuel Fernández-Delgado, Eva Cernadas, Senén Barro, and Dinani Amorim. 2014. Do we need hundreds of classifiers to solve real world classification problems. *Journal of Machine Learning Research (JMLR)*, 15(1):3133–3181.

Lucie Flekova and Iryna Gurevych. 2013. Can we hide in the web? large scale simultaneous age and gender author profiling in social media. In *Proceedings of the Conference and Labs of the Evaluation Forum (CLEF)*.

Aparna Garimella and Rada Mihalcea. 2016. Zooming in on gender differences in social media. In *Proceedings of the COLING Workshop on Computational Modeling of People's Opinions, Personality, and Emotions in Social Media (PEOPLES)*, pages 1–10.

Sumit Goswami, Sudeshna Sarkar, and Mayur Rustagi. 2009. Stylometric analysis of bloggers' age and gender. In *Proceedings of the International Conference on Web and Social Media (ICWSM)*.

Suparna Goswami, Felix Köbler, Jan Marco Leimeister, and Helmut Krcmar. 2010. Using online social networking to enhance social connectedness and social support for the elderly. In *Proceedings of the International Conference on Information Systems (ICIS)*. Association for Information Systems.

Foad Hamidi, Morgan Klaus Scheuerman, and Stacy M Branham. 2018. Gender recognition or gender reductionism?: The social implications of embedded gender recognition systems. In *Proceedings of the Conference on Human Factors in Computing Systems (CHI)*, page 8. ACM.

Susan C Herring. 2003. Gender and power in on-line communication. *The Handbook of Language and Gender*, pages 202–228.

Heather J Hether, Sheila T Murphy, and Thomas W Valente. 2016. A social network analysis of supportive interactions on prenatal sites. *Digital Health*, 2:2055207616628700.

Sepp Hochreiter and Jürgen Schmidhuber. 1997. Long short-term memory. *Neural Computation*, 9(8):1735–1780.

Matthew Honnibal and Mark Johnson. 2015. An improved non-monotonic transition system for dependency parsing. In *Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 1373–1378.

Dirk Hovy and Anders Søgaard. 2015. Tagging performance correlates with author age. In *Proceedings of the Annual Meeting of the Assocation for Computational Linguistics (ACL)*, pages 483–488.

Eduard Hovy. 1987. Generating natural language under pragmatic constraints. *Journal of Pragmatics*, 11(6):689–719.

Sergey Ioffe and Christian Szegedy. 2015. Batch normalization: Accelerating deep network training by reducing internal covariate shift. In *Proceedings of the International Conference on Machine Learning (ICML)*, pages 448–456.

Aaron Jaech and Mari Ostendorf. 2015. What your username says about you. In *Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 2032–2037.

Hamed Khanpour, Cornelia Caragea, and Prakhar Biyani. 2018. Identifying emotional support in online health communities. In *Proceedings of the Conference on Artificial Intelligence (AAAI)*.

Sara Kiesler, Jane Siegel, and Timothy W McGuire. 1984. Social psychological aspects of computer-mediated communication. *American Psychologist*, 39(10):1123.

Diederik P Kingma and Jimmy Ba. 2014. Adam: A method for stochastic optimization. In *Proceedings of the International Conference on Learning Representations (ICLR)*.

Rebecca Knowles, Josh Carroll, and Mark Dredze. 2016. Demographer: Extremely simple name demographics. In *EMNLP Workshop on NLP and Computational Social Science*, pages 108–113. Association for Computational Linguistics.

Victor Kuperman, Hans Stadthagen-Gonzalez, and Marc Brysbaert. 2012. Age-of-acquisition ratings for 30,000 english words. *Behavior Research Methods*, 44(4):978–990.

Shyong Tony K Lam, Anuradha Uduwage, Zhenhua Dong, Shilad Sen, David R Musicant, Loren Terveen, and John Riedl. 2011. Wp: clubhouse?: An exploration of wikipedia's gender imbalance. In *Proceedings of the 7th International Symposium on Wikis and Open Collaboration (WikiSym)*, pages 1–10. ACM.

Brian N Larson. 2017. Gender as a variable in natural-language processing: Ethical considerations. In *Proceedings of the First Workshop on Ethics in Natural Language Processing*.

Martin Lea and Russell Spears. 1991. Computer-mediated communication, de-individuation and group decision-making. *International Journal of Man-Machine Studies*, 34(2):283–301.

Jia Li, Xuan Liu, and Min Sun. 2018. Research on gender differences in online health communities. *International Journal of Medical Informatics*.

Siyue Li and Bo Feng. 2014. What to say to an online support-seeker? the influence of others' responses and support-seekers' replies. *Human Communication Research*, 41(3):303–326.

Bing Liu, Minqing Hu, and Junsheng Cheng. 2005. Opinion observer: analyzing and comparing opinions on the web. In *Proceedings of the International Conference on World Wide Web (WWW)*, pages 342–351. ACM.

Wendy Liu and Derek Ruths. 2013. What's in a name? using first names as features for gender inference in twitter. In *AAAI Spring Symposium: Analyzing Microtext*, pages 10–16.

Ido Liviatan, Yaacov Trope, and Nira Liberman. 2008. Interpersonal similarity as a social distance dimension: Implications for perception of others' actions. *Journal of Experimental Social Psychology*, 44(5):1256–1269.

Erina L MacGeorge, Bo Feng, and Brant R Burleson. 2011. Supportive communication. *Handbook of Interpersonal Communication*, pages 317–354.

Gabriel Magno and Ingmar Weber. 2014. International gender differences and gaps in online social networks. In *Proceedings of the International Conference on Social Informatics (SocInfo)*, pages 121–138. Springer.

Alice Marwick. 2013. Online identity. *A companion to new media dynamics*.

Michael McCandless. 2010. Accuracy and performance of Google's compact language detector. http://blog.mikemccandless.com/2011/10/accuracy-and-performance-of-googles.html.

Jessica Megarry. 2014. Online incivility or sexual harassment? conceptualising women's experiences in the digital age. In *Women's Studies International Forum*, volume 47, pages 46–55. Elsevier.

Kristin D Mickelson, Vicki S Helgeson, and Erica Weiner. 1995. Gender effects on social support provision and receipt. *Personal Relationships*, 2(3):211–224.

Tomas Mikolov, Ilya Sutskever, Kai Chen, Greg S Corrado, and Jeff Dean. 2013. Distributed representations of words and phrases and their compositionality. In *Proceedings of the International Conference on Neural Information Processing Systems (NIPS)*, pages 3111–3119.

Saif M. Mohammad and Peter D. Turney. 2013. Crowdsourcing a word-emotion association lexicon. *Artificial Intelligence*, 29(3):436–465.

Ellie Pavlick and Joel Tetreault. 2016. An empirical analysis of formality in online communication. *Transactions of the Association of Computational Linguistics (TACL)*, 4(1):61–74.

Fabian Pedregosa, Gaël Varoquaux, Alexandre Gramfort, Vincent Michel, Bertrand Thirion, Olivier Grisel, Mathieu Blondel, Peter Prettenhofer, Ron Weiss, Vincent Dubourg, et al. 2011. Scikit-learn: Machine learning in python. *Journal of Machine Learning Research (JMLR)*, 12:2825–2830.

James W Pennebaker and Lori D Stone. 2003. Words of wisdom: language use over the life span. *Journal of Personality and Social Psychology*, 85(2):291.

Tom Postmes and Russell Spears. 2002. Behavior online: Does anonymous computer communication reduce gender inequality? *Personality and Social Psychology Bulletin*, 28(8):1073–1083.

Tom Postmes, Russell Spears, and Martin Lea. 1998. Breaching or building social boundaries? side-effects of computer-mediated communication. *Communication Research*, 25(6):689–715.

Randolph Quirk, Sidney Greenbaum, Geoffrey Leech, and Jan Svartvik. 1985. A comprehensive grammar of the english language. *London and New York: Longman*.

Stephen A Rains. 2016. Language style matching as a predictor of perceived social support in computer-mediated interaction among individuals coping with illness. *Communication Research*, 43(5):694–712.

Sravana Reddy and Kevin Knight. 2016. Obfuscating gender in social media writing. In *Proceedings of Workshop on Natural Language Processing and Computational Social Science*, pages 17–26.

Tyler Joseph Schnoebelen. 2012. *Emotions are relational: positioning and the use of affective linguistic resources*. Ph.D. thesis, Stanford University.

Chris Segrin. 2003. Age moderates the relationship between social support and psychosocial problems. *Human Communication Research*, 29(3):317–342.

Sally A Shumaker and Arlene Brownell. 1984. Toward a theory of social support: Closing conceptual gaps. *Journal of Social Issues*, 40(4):11–36.

Robert J Sigley. 1997. Text categories and where you can stick them: a crude formality index. *International Journal of Corpus Linguistics*, 2(2):199–237.

Russell Spears and Martin Lea. 1992. Social influence and the influence of the "social" in computer-mediated communication. In Martin Lea, editor, *Contexts of Computer-Mediated Communication*, pages 30–65). Harvester Wheatsheaf.

Erin L Spottswood, Joseph B Walther, Amanda J Holmstrom, and Nicole B Ellison. 2013. Person-centered emotional support and gender attributions in computer-mediated communication. *Human Communication Research*, 39(3):295–316.

Cong Tang, Keith Ross, Nitesh Saxena, and Ruichuan Chen. 2011. What's in a name: A study of names, gender inference, and gender behavior in facebook. In *International Conference on Database Systems for Advanced Applications*, pages 344–356. Springer.

Simone Teufel. 2000. *Argumentative zoning: Information extraction from scientific text*. Ph.D. thesis, University of Ediburgh.

Alan Vaux. 1985. Variations in social support associated with gender, ethnicity, and age. *Journal of Social Issues*, 41(1):89–110.

Rob Voigt, David Jurgens, Vinodkumar Prabhakaran, Dan Jurafsky, and Yulia Tsvetkov. 2018. Rtgender: A corpus for studying differential responses to gender. In *Proceedings of the Language Resources and Evaluation Conference (LREC)*.

Svitlana Volkova, Yoram Bachrach, Michael Armstrong, and Vijay Sharma. 2015. Inferring latent user properties from texts published in social media. In *Proceedings of the AAAI Conference on Artificial Intelligence (AAAI)*, pages 4296–4297.

Sida Wang and Christopher D Manning. 2012. Baselines and bigrams: Simple, good sentiment and topic classification. In *Proceedings of the Annual Meeting of the Association for Computational Linguistics (ACL)*, pages 90–94. Association for Computational Linguistics.

Yi-Chia Wang, Robert Kraut, and John M Levine. 2012. To stay or leave?: the relationship of emotional and informational support to commitment in online health support groups. In *Proceedings of the Conference on Computer Supported Cooperative Work (CSCW)*, pages 833–842. ACM.

Zuoming Wang, Joseph B Walther, and Jeffrey T Hancock. 2009. Social identification and interpersonal communication in computer-mediated communication: What you do versus who you are in virtual groups. *Human Communication Research*, 35(1):59–85.

Barry Wellman and Scot Wortley. 1990. Different strokes from different folks: Community ties and social support. *American Journal of Sociology*, 96(3):558–588.

Kevin Wright. 2002a. Motives for communication within on-line support groups and antecedents for interpersonal use. *Communication Research Reports*, 19(1):89–98.

Kevin Wright. 2002b. Social support within an online cancer community: An assessment of emotional support, perceptions of advantages and disadvantages, and motives for using the community from a communication perspective. *Journal of Applied Communication Research*, 30(3):195–209.

Kevin B Wright, Sally B Bell, Kevin B Wright, and Sally B Bell. 2003. Health-related support groups on the internet: Linking empirical findings to social support and computer-mediated communication theory. *Journal of Health Psychology*, 8(1):39–54.