Linguistic synchrony predicts the immediate and lasting impact of online peer-to-peer emotional support

[1]Doré, BP & [2]Morris, RR

Affiliations:
[1] Annenberg School for Communication, University of Pennsylvania, Philadelphia, PA
[2] Koko, New York, NY

Correspondence to: brucedore@gmail.com

Abstract WC: 138
Intro + Discussion + Acknowledgments WC: 750
Methods + Results WC: 2938
Figures: 2

Abstract

Emotional support is critical to well-being, but the factors that determine whether support attempts succeed or fail are incompletely understood. Using data from more than one million support interactions enacted within an online environment, we show that emotional support attempts are more effective when there is synchrony in the behavior of support providers and recipients reflective of shared psychological understanding. Benefits of synchrony in language used and semantic content conveyed were apparent in immediate measures of support impact (recipient ratings of support effectiveness and expressions of gratitude), as well as delayed measures of lasting change in the emotional impact of stressful life situations (recipient ratings of emotional recovery made at a one-hour delay). These findings identify linguistic synchrony as a process underlying successful emotional support and provide direction for future work investigating support processes enacted via linguistic behaviors.

When confronted with distressing experiences, we often reach to others for support in managing our emotions. Unfortunately, it can be difficult to provide such support effectively, and the processes that promote emotional support efficacy are incompletely understood (Gleason et al., 2008; Goldsmith, 2004).

Prior work has described effective support in terms of multifaceted constructs like the responsiveness and sensitivity of the support provider (Goldsmith, 2004; Reis & Gable, 2015). In the present study, we sought to provide mechanistic insight into emotional support exchanges by leveraging computational analysis of text-based social interactions. Drawing from prior theorizing, we reasoned that emotional support would be more impactful when support providers and recipients show coordination of their linguistic behavior that reflects shared psychological understanding (Giles, 2016; Ireland et al., 2011; Reis & Gable, 2015). In particular, we hypothesized that synchrony in words used and semantic meaning conveyed can influence how emotional support attempts are initially received, and whether they evoke durable change in emotion.

To test this hypothesis, we analyzed instances of emotional support that occurred within an online social network (Morris, Schueller, & Picard, 2015). Within this network, users anonymously posted descriptions of stressful life experiences and received supportive responses from other users. We considered several outcomes of emotional support, including recipients' ratings of support effectiveness, expressions of gratitude (thank-you notes), and ratings of lasting emotional recovery. With these data, we asked two key questions: 1) whether recipient-provider synchrony in textual content, linguistic style, kinds of emotions expressed, and latent meaning conveyed predicted these support outcomes, and 2) how synchrony effects compared to those of another fundamental property of emotional support – the positivity of the language it uses.

Method

Participants and Design

Participants in this study were users of an online application, called *Koko* (previously *Panoply*), that facilitates text-based emotional support interactions within an anonymous social network (Morris, Schueller, & Picard, 2015). We analyzed data from every user who interacted with this application between June 1, 2016 and June 20, 2017, a total of 169,376 unique users posting about 361,139 unique stressful situations and receiving 1,161,360 messages of support in response. Users could learn about this application in a variety of ways (e.g., online advertisements, news articles, referrals from social media applications, and word-of-mouth) and interface with it through several social media channels (e.g., Facebook Messenger, Twitter, Telegram Messenger, Kik Messenger, and a standalone iPhone application). Since the application is anonymous, we did not collect demographic information like age or gender. However, users of these kinds of social media channels tend to be younger than the general population (Pew Research Center, 2017). This dataset was originally collected for internal evaluation and improvement of the application. Analysis of this pre-existing dataset was deemed exempt from review by the University of Pennsylvania Institutional Review Board.

Across different channels for entry, the experience of interacting with the application included the same core elements: initial onboarding, training in how to use the application, anonymously posting about life stressors to solicit emotional support from other users, and composing and sending supportive messages in response to the posts of others (see Supplementary Materials). After onboarding and training, users were invited to write their first stressor post. Specifically, they were asked to describe a life experience that is a current source of stress, and were further prompted to describe their negative thoughts about this experience (see Figure 1). After submitting this stressor post, users began to receive supportive messages from other users, typically receiving messages of support from 3 or 4 other users starting a few minutes after posting (median delay from post to a response was 6.2 minutes, with

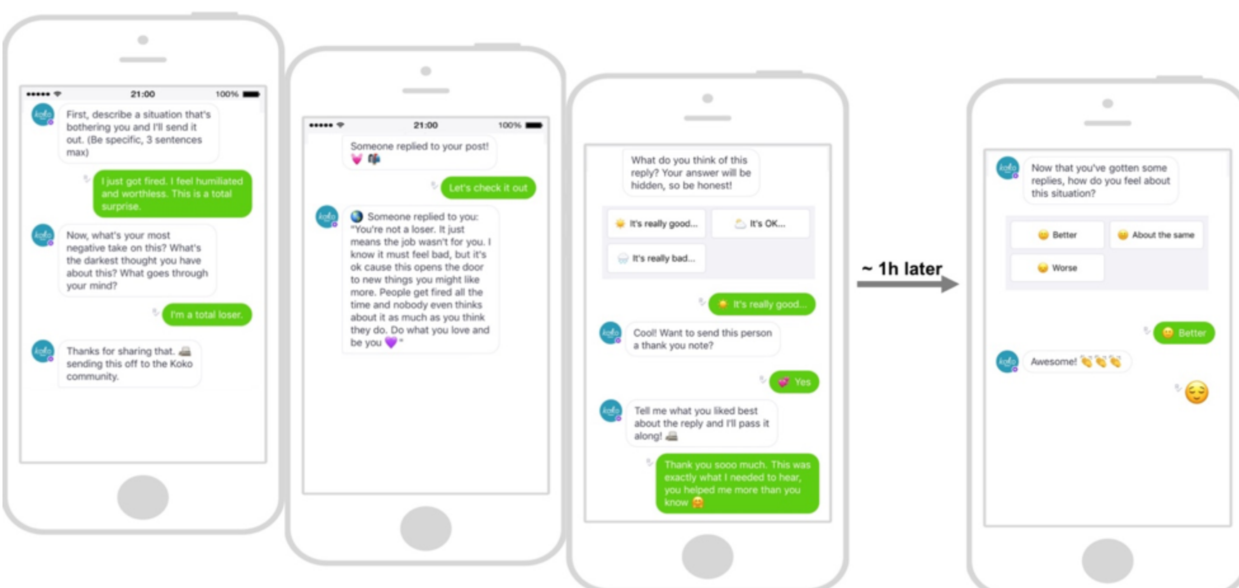

Figure 1. An example exchange between two users, consisting of **A)** a recipient posting thoughts and feelings surrounding a stressful life situation, **B)** a response from another user providing emotional support, and **C)** an immediate rating of support effectiveness and an opportunity to send a thank-you note. **D)** After a delay of at least 30 minutes and typically about one hour, the support recipient rated their current feelings about the the situation they posted about.

interquartile range of 3.8 minutes to 16.1 minutes). Users providing support were given minimal training in how to provide supportive responses, and ask to describe a more positive take on the described post. After submitting their first stressor post, users were able to 1) read descriptions of stressful experiences posted by other users, 2) compose and send supportive responses to these posts, and 3) submit additional stressor posts describing other stressful experiences in their lives. As is typical for counts of ecological behaviors, individual differences in the volume of platform behaviors per user could be well described with negative binomial distributions, for both stressor posts (location parameter $\mu = 1.1$ posts, dispersion parameter $\theta = 1.9$), and for support responses (location parameter $\mu = 2.3$ responses, dispersion parameter $\theta = 13.9$). Immediately after receiving a support response, users were asked to rate the response's effectiveness and were given the option of sending a brief thank you message (median length 11 words, interquartile range 5 words to 20 words). After a delay of at least 30 minutes after support receipt, users were additionally asked to provide a rating about how they currently felt about the stressful life experience they posted about.

Measures and Analyses

Predictor variables. We derived predictor variables that were based on similarity between recipients' stressor posts (i.e., the description of the stressful life experience, combined with the additional description of their negative thoughts about the experience) and providers' supportive responses (i.e., the texts sent in reply to the posts that aimed to provide emotional support). Similarity computations and all subsequent analyses were performed in R (version 3.4.0). Our selection of predictor variables was informed by a model positing that synchrony in textual content, style words, emotion words, and latent semantic content reflects a support provider who is able to: use language from the recipient's post, reference relevant actors or objects of the post using the relevant function words, reference or re-express emotional states expressed by the recipient, and in a broader sense speak to the semantic content conveyed by the recipient (i.e., semantic content that is not reflected in function words or emotion words).

We defined surface-level textual similarity as the opposite of the Levenshtein distance, which is the minimum number of single-character deletions, insertions, or substitutions required to change one text into another (Levenshtein, 1966). The greater the Levenshtein distance, the more different two documents are in terms of the text they use.

We defined synchrony in linguistic style (typically called *language style matching*) as similar use of the *function words* from the Linguistic Inquiry and Word Count (LIWC) dictionary – negations, quantifiers, conjunctions, adverbs, auxiliary verbs, prepositions, articles, personal pronouns, and impersonal pronouns (Neiderhoffer & Pennebaker, 2002; Ireland et al., 2011; Pennebaker, Booth, & Francis, 2007). Following previous research, we defined synchrony in linguistic style using formulas of the form:

$$Synchrony_{preps} = 1 - [(\left| preps_{post} - preps_{response} \right|)/(preps_{post} + preps_{response} + 0.0001)]$$

In this formula, $preps_{post}$ is the percentage of prepositions used by the support recipient in the stressor post and $preps_{response}$ is the percentage of prepositions used by the support provider within the supportive response. In the denominator, 0.0001 is added to prevent empty sets. The nine scores for each function word category were averaged to yield a composite score bounded by 0 and 1; a higher number represents greater synchrony in linguistic style (i.e. similarity in use of function words) across a recipients' stressor posts and a providers' supportive response. Prior work has shown that language style matching can predict social outcomes, like relationship initiation and group cohesion (Gonzales,Hancock, & Pennebaker, 2010; Ireland et al., 2011).

We also considered synchrony in the texts' overall emotional character (i.e., the specific kinds of emotions expressed). We defined the texts' emotional character using a lexicon-based algorithm that estimates expression of eight categories of emotion -- joy, trust, fear, surprise, sadness, disgust, anger, and anticipation -- , as well as overall positive and negative valence, across 14,182 English words (Mohammad & Turney, 2013). We used this lexicon because it provides a broader assessment of emotional character than other commonly used sentiment analysis tools (i.e., it includes more categories of emotion and a larger lexicon than alternatives) and it has been shown to perform well in capturing emotional content of social media texts (Ribeiro et al., 2016). We defined synchrony in emotional character in a manner analogous to that used for linguistic style synchrony (Neiderhoffer & Pennebaker, 2002; Ireland et al., 2011), with the only difference being that the computation was made on the basis of categories of emotion words (Mohammad & Turney, 2013) rather than categories of function words. That is, we used formulas with the form:

$$Synchrony_{fear} = 1 - [(\left| fear_{post} - fear_{response} \right|)/(fear_{post} + preps_{response} + 0.0001)]$$

Here $fear_{post}$ is the percentage of words associated with the fear category used by the recipient in the stressor post, and $fear_{response}$ is the percentage of words associated with the feat category used by the provider in the supportive response. As with synchrony in function word use, the eight synchrony scores for each category of emotion were averaged to yield a composite score bounded by 0 and 1; a higher number represent greater synchrony in emotional character (though not necessarily in the specific words used) across a recipients' stressor post and a providers' supportive response. We also used the same lexicon to estimate the overall valence – defined as the score for positive sentiment minus the score for negative sentiment – of stressor posts, supportive responses, and thank you notes.

Finally, we considered similarity in latent semantic content – that is, similarity in the kinds of underlying topics that are addressed in the stressor posts and support responses, despite potential differences in the words and phrases used. In order to do this, we used latent semantic analysis (Landauer, 2007). Latent semantic analysis (LSA) provides an algorithmic estimate of document similarity via a two-step procedure. In the first step, a large set of words (here, over two billion English words used in online writing) are reduced to a lower-rank set of latent semantic vectors based on word co-occurrence across a large set of documents (here, a corpus of online documents). In a next step, latent semantic similarity is estimated by computing the cosine similarity between the latent semantic vectors expressed across two texts (here a stressor post and its corresponding support response). The cosine similarity measures the angle between these two semantic vectors, capturing the idea that texts that are similar in meaning should exist close to each other within a multidimensional semantic space. In summary, this technique represents the meaning of a particular English word as reflecting the contexts in which the word tends to appear, and further represents the meaning of a particular document as reflecting the meanings of all the words it contains. In this sense, the basic idea behind latent semantic analysis is that documents that use similar kinds of words at similar frequency are semantically related. Prior work supports the validity of latent semantic similarity as a measure of semantic relatedness; for example, LSA can predict human ratings of semantic relatedness, approach human levels of accuracy in assessing essay content, and achieve a passing grade in a college multiple choice text after being trained on a relevant textbook (reviewed in Landauer, 2007). For a discussion of how this method is sensitive to the presence of grammatical negations, and for examples of post-response pairs that were estimated to be low, medium, and high in latent semantic similarity (as well as emotional content similarity and linguistic style similarity), see Supplementary Materials.

Outcome variables. Outcome variables were based on the ratings and behavioral responses made by support recipients (see Figure 1). Three of these variables were collected immediately after receipt of a support response: immediate ratings of support effectiveness (*"What do you think of this reply? Your answer will be hidden, so be honest!"* -1: *it's really bad*, 0: *it's ok…*, +1: *it's really good)*, whether the support recipient decided to send a thank you not to the support provider (0: did not send, 1: did send), and the expressed valence of the thank you note (mean valence = +1.27, standard deviation = 2.41). The fourth outcome variable of interest was ratings of stressor-specific emotional change collected after a delay. Ratings of stressor-specific emotional change ratings were sent 30 minutes (the median time elapsed from original post to rating emotional recovery was 63min, interquartile range of 40min to 8h) after beginning to receive support responses (“*Now that you've gotten some replies, how do you feel about your situation?”* -1: *worse*, 0: *about the same*, +1: *better*). When computing synchrony for the purpose of predicting these delayed ratings we calculated the average degree of synchrony for all the support responses sent in response to a given post – the average degree of synchrony in messages directed to that post (for a description of variability in post volume, see Supplementary Materials).

Modelling. We used boxplot and wordclouds for initial visualizations of the content of stressor posts and response texts. Word clouds were generated using a random sample of 100,000 unique stressor posts and 100,000 unique support responses. To minimize overfitting in our primary analyses, we randomly split the data into an exploratory sample of 33,875 participants (20% of the full dataset) for initial visualization and model building, and a confirmatory sample of 135,501 participants (80% of the full dataset) used to evaluate the fit of regularized regression models and thereby support statistical inference. Specifically, we fit generalized additive mixed models (GAMMs), a technique from machine learning, to estimate relationships between synchrony and recipient ratings and behavioral responses (Wood, 2017). In the GAMM framework, an outcome variable (e.g., ratings of support effectiveness) varies as an unknown smooth function of a predictor variable (e.g., semantic similarity of stressor posts and support responses), and this function is represented using regression splines (i.e., piecewise polynomial fits connected by knots). Model form and smoothness are not user-specified, as when selecting a linear, quadratic, or n-degree polynomial fit in regression, but rather estimated from the data via a fitting procedure in which an optimal smoothness is selected by penalized likelihood metrics that approximate out-of-sample predictive accuracy (Wood, 2017). This procedure yields a regularized estimate of the population-level relationship with credibility intervals reflecting differences in form this function could plausibly take in light of the observed data. It also yields an overall *p* value reflecting compatibility of the data with a null model (i.e., a flat line). We set the *k* parameter (upper limit on effective degrees of freedom) for all smooth functions to 8, allowing the estimated functions a high but not extreme degree of flexibility (Wood, 2017). Because different measures of synchrony were weakly or moderately correlated, models included all four kinds of synchrony as simultaneous predictors; that is, effects of a given kind of synchrony were estimated after adjusting for effects of the other kinds of synchrony. For models predicting stressor-specific emotional recovery ratings, we computed similarity for each response with its corresponding post, and then averaged these similarity values to yield a single number reflecting the average degree of synchrony in the support responses received for a particular stressor post.

Results

In an initial step, we used text analytic methods to provide insight into the content of the posts describing thoughts about life stressors and responses providing emotional support. Using a lexicon-based algorithm to estimate emotional expression (Mohammad & Turney, 2016), we found that posts describing stressors were negative in valence and expressed mostly sadness, anticipation, fear, and anger categories of emotion consistent with instructions to describe thoughts and feelings surrounding a stressful experience. Responses providing emotional support were more positive in valence and expressed more joy and trust categories of emotion, consistent with instructions to provide a more positive and supportive take on the experience (see Supplementary Figure 1). To shed light on similarities and differences in word use across stressor posts versus support responses, we visualized word use across stressor posts and support responses (see Supplementary Figures S2 and S3). These visualizations suggested that posts and responses were similar in generally referencing social relationships and psychological states, but differed in specific use of certain function words (e.g., pronouns), emotion words, and words referencing specific social relationships.

Our primary research question was whether synchrony between stressor posts and supportive responses was related to emotional support effectiveness. We first considered synchrony in the actual textual characters used. We defined this as the opposite of the Levenshtein distance – the number of single-character edits needed to transform one text into another. We asked whether surface-level textual synchrony predicted immediate ratings of support effectiveness, whether the recipient sent an expression of gratitude (‘thank you note’), and the valence of language used within expressions of gratitude. Across these three outcomes, support responses that were highly asynchronous (showed very little overlap in text used) or highly synchronous (repeated much of the stressor post in a close to verbatim manner) with their corresponding stressor post were less impactful than those showing a moderate degree of surface-level textual similarity (see Figure 2).

We next considered synchrony in linguistic style, defined as similar use of function words like pronouns and adverbs (Ireland et al., 2011), and synchrony in emotional content, defined as similar use of words from different emotion categories, like fear versus sadness (Mohammad & Turney, 2013). For both linguistic style synchrony and emotional content synchrony, support responses showing greater synchrony were rated more effective, were more likely to elicit an expression of gratitude, and elicited more positive language within gratitude expressions. However, these relationships showed clear nonlinearity in a manner suggesting diminishing additional benefits of high (versus moderate) synchrony in linguistic style (see Figure 2).

We next turned to post-response synchrony in latent semantic content, defined using latent semantic analysis (Landauer, 2007). In contrast to linguistic style synchrony and emotional content synchrony, latent semantic synchrony robustly predicted immediate support outcomes with a form suggesting an exponential relationship, indicating that synchrony in semantic meaning was an especially powerful source of emotional support efficacy relative to the other kinds of synchrony we examined.

Having identified a role for synchrony in predicting immediate support outcomes, we next asked if support synchrony was related to emotional recovery across a longer timescale. Within the application, recipients rated whether they felt better, the same, or worse about the stressor they described about one hour after receiving supportive responses (median 63min, interquartile range 40min to 8hr). As shown in the bottom row of Figure 2, these ratings of lasting recovery were related to each metric of synchrony. In particular, latent semantic synchrony showed a clear positive relationship with lasting emotional recovery, suggesting that support that was attuned to the meaning of the language used by recipients was ultimately more impactful. However, although synchronous linguistic style and emotional content predicted higher immediate ratings of effectiveness, non-monotonic relationships with these delayed ratings suggested that moderate (rather than high or low) synchrony on these dimensions was more optimal for evoking lasting emotional change. Further, these effects were unchanged when additionally controlling for any smooth effect of time to rating (see Supplementary Materials).

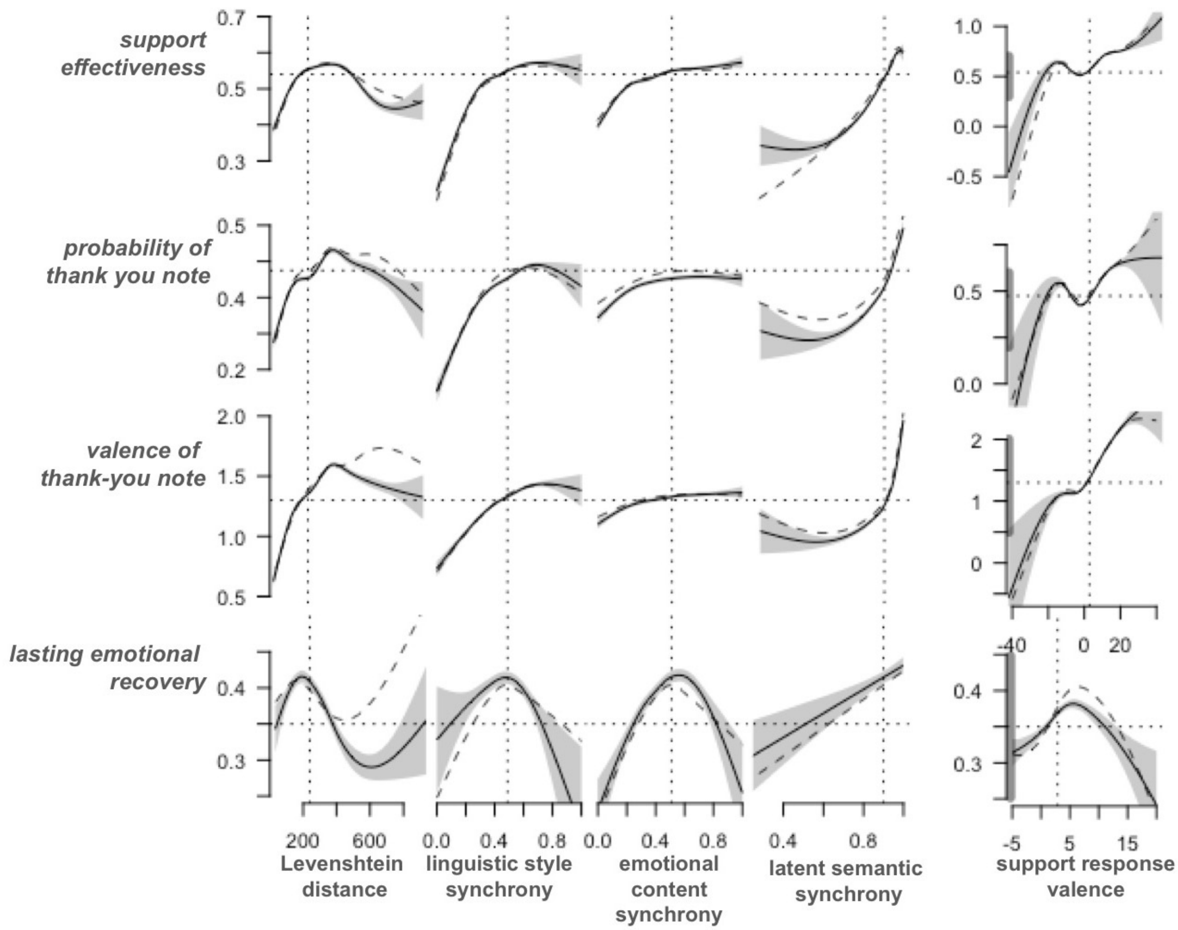

Figure 2. Functions relating textual, linguistic style, emotional content and latent semantic synchrony, as well as overall support response valence, to immediate ratings of emotional support effectiveness, whether a thank you note was sent, valence of thank you note language, and lasting emotional recovery. Notes: solid curves with 95% credibility intervals reflect models fit to confirmatory subsample (80% of the data), all significant at $p<10^{-8}$; dashed curves reflect models fit to exploratory subsample (20% of the data); dotted horizontal and vertical lines give reference to the mean of each variable.

Finally, we contrasted support synchrony with another fundamental property of emotional support – the overall valence of the language it uses. As shown in Figure 2 (rightmost column), there were large differences in the outcomes seen for highly negative versus neutral support responses and smaller differences between neutral versus highly positive support responses. Gauging predictive effects of support synchrony versus valence, two findings stood out. First, effects of synchrony were somewhat smaller than effects of valence for immediate support outcomes (i.e., effectiveness ratings and thank-you notes). Second, synchrony and valence were different in their implications for lasting emotional recovery. Latent semantic synchrony linearly predicted higher lasting recovery, but support valence showed a non-monotonic relationship such that lasting recovery was seen most for support with moderately (but not extremely) positive language.

Discussion

Overall, these results suggest that emotional support is more effective when there is synchrony in the behavior of support providers and recipients, especially synchrony reflective of shared understanding. For textual content, linguistic style, and emotional content, a moderate degree of synchrony predicted beneficial outcomes. For latent semantic content, higher than average levels of synchrony predicted greater immediate and lasting impact of emotional support.

These findings have implications for theories of the processes that underlie successful regulation of emotion within social interactions (Goldsmith, 2004; Reis & Gable 2015; Zaki & Williams, 2013). First, they suggest that support is most effective when it speaks to the meaning of what is communicated by people seeking help.

Second, they suggest that support is more effective when it uses language that is neither overly convergent nor overly discrepant with the language used by people seeking help. Support that was highly synchronous to the recipient narrative in linguistic style and emotional content was initially well-received but ultimately less effective in changing emotion, suggesting that providers can over-synchronize along these dimensions. This is consistent with the notion that synchrony in function words reflects a kind of active conversational engagement, but suggests that an excessive degree of this engagement may be suboptimal for emotional support (Niederhoffer & Pennebaker, 2002). Overall, these data suggest that similarity in style, emotional content, and semantic content index distinct psychological processes that together scaffold effective support. In light of these data, we propose that effective emotional support derives from a shared understanding that is apparent in linguistic behavior, and that may emerge from recipient characteristics like expressiveness, provider characteristics like empathic ability, and recipient-provider effects like similarity in life history or personality (Goldsmith, 2004; Cavallo, Zee, & Higgins, 2016; Giles, 2016).

In parallel, our results speak to a changing media landscape where support increasingly unfolds online. Some data suggest that computer-mediated support can help people overcome challenges associated with seeking and providing support face-to-face (Pentina & Zhang, 2017). However, most online spaces are not designed to promote well-being, and some evidence suggests they can be actively harmful (Verduyn et al., 2017). Future work could evaluate how synchronous and non-synchronous elements of social interactions (e.g., offering new perspective, or strong but helpful criticism) work together to promote well-being and social understanding, examining implications for the design of online environments. Further, future work could improve on the single-item measures of emotion we used here (for which internal consistency is undefined) by using multi-item measures, and could also ask about serial position effects within a stream of emotional support responses.

When people around us are struggling, we often do what we can to help them manage their emotions. However, attempts to provide such help can fail to have their intended impact. We suggest that linguistic synchrony predicts emotional support efficacy, especially synchrony that reflects a shared understanding of the meaning conveyed within a social exchange. We hope that future studies will expand on our approach to further unpack mechanisms that underlie the ability to collectively navigate life's emotional challenges.

References

Cavallo, J. V., Zee, K. S., & Higgins, E. T. (2016). Giving the Help That Is Needed: How Regulatory Mode Impacts Social Support. *Personality & Social Psychology Bulleti*n, 42(8), 1111–1128.

Giles, H. (2016). Communication Accommodation Theory. In The International Encyclopedia of Communication Theory and Philosophy. John Wiley & Sons, Inc.

Gleason, M. E. J., Iida, M., Shrout, P. E., & Bolger, N. (2008). Receiving support as a mixed blessing: evidence for dual effects of support on psychological outcomes. *Journal of Personality and Social Psychology*, 94(5), 824–838.

Goldsmith, D. J. (2004). Communicating social support. New York, NY: Cambridge University Press.

Gonzales, A. L., Hancock, J. T., & Pennebaker, J. W. (2010). Language style matching as a predictor of social dynamics in small groups. *Communication Research,* 37(1), 3-19.

Ireland, M. E., Slatcher, R. B., Eastwick, P. W., Scissors, L. E., Finkel, E. J., & Pennebaker, J. W. (2011). Language style matching predicts relationship initiation and stability. *Psychological Science*, 22(1), 39–44.

Landauer, T. K. (2007) LSA as a theory of meaning. In T. K.Landauer, D. S.McNamara, S.Dennis, & W.Kintsch (Eds.), *Handbook of latent semantic analysis* (pp. 3–34). Mahwah, NJ: Erlbaum.

Levenshtein, V. I. (1966). Binary codes capable of correcting deletions, insertions, and reversals. *Soviet physics doklady*,10, 707–710.

Mohammad, S. M., & Turney, P. D. (2013). Crowdsourcing a word–emotion association lexicon. *Computational Intelligence. An International Journal*, 29(3), 436–465.

Morris, R. R., Schueller, S. M., & Picard, R. W. (2015). Efficacy of a web-based, crowdsourced peer-to-peer cognitive reappraisal platform for depression: Randomized controlled trial. *Journal of medical Internet research*, 17(3).

Niederhoffer, K. G., & Pennebaker, J. W. (2002). Linguistic style matching in social interaction. Journal of Language and Social Psychology, 21(4), 337-360.

Pennebaker, J. W., Booth, R. J., & Francis, M. E. (2007). LIWC2007: Linguistic inquiry and word count. Austin, Texas: liwc.net.

Pentina, I., & Zhang, L. (2017). Effects of social support and personality on emotional disclosure on Facebook and in real life. *Behaviour & Information Technology,* 36(5), 484–492.

Pew Research Center (2017). Social Media Fact Sheet. Retrieved from http://www.pewinternet.org/fact-sheet/social-media/

Reis, H. T., & Gable, S. L. (2015). Responsiveness. *Current Opinion in Psychology*, 1, 67-71.

Ribeiro, F. N., Araújo, M., Gonçalves, P., Gonçalves, M. A., & Benevenuto, F. (2016). SentiBench - a benchmark comparison of state-of-the-practice sentiment analysis methods. *EPJ Data Science*, 5(1), 23.

Verduyn, P., Ybarra, O., Résibois, M., Jonides, J., & Kross, E. (2017). Do Social Network Sites Enhance or Undermine Subjective Well-Being? A Critical Review. *Social Issues and Policy Review*, 11(1), 274-302.

Wood, S. N. (2017). Generalized Additive Models: An Introduction with R, Second Edition. CRC Press.

Zaki, J., & Williams, W. C. (2013). Interpersonal emotion regulation. *Emotion*, 13(5), 803–810.

Acknowledgments

We thank Yuriy Blokhin for technical support, Travis Riddle and Katherine Zee for helpful discussion, and users of the Koko app for making this work possible.