How to be Helpful on Online Support Forums?

Zhilin Wang
University of Washington
zhilinw@uw.edu

Pablo Torres
University of Cambridge
pelt2@cam.ac.uk

Abstract

Internet forums such as Reddit offer people a platform to ask for advice when they encounter various issues at work, school or in relationships. Telling helpful comments apart from unhelpful comments to these advice-seeking posts can help people and dialogue agents to become more helpful in offering advice. We propose a dataset that contains both helpful and unhelpful comments in response to such requests. We then relate helpfulness to the closely related construct of empathy. Finally, we are the first study to analyze the language features that are associated with helpful and unhelpful comments.

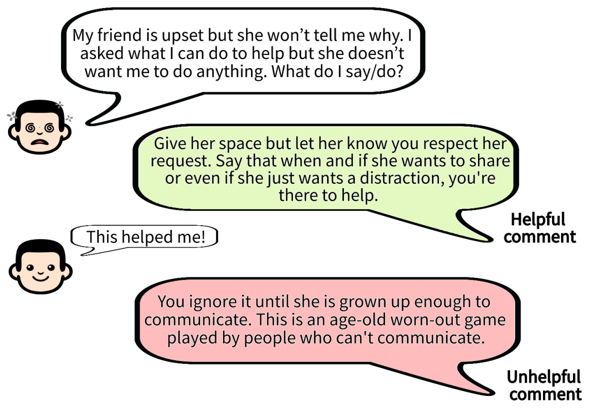

Figure 1: Examples of helpful and un-helpful comments to a help-seeking post

1 Introduction

When people encounter issues in their lives (such as problems with family and friends, difficulties at school/work as well as troubles in pursuing one's interests and hobbies), many seek for advice in order to solve these problems. Some ask for such advice on internet forums, such as the r/Advice subreddit[1]. Other users can then comment on these posts to attempt to help the post authors.

While many users can actively offer help, not all of them will be seen as helpful by the user asking for advice. Examples of a helpful and an unhelpful comment are presented in Figure 1 to show their contrast. In order to support people and dialogue agents to be more effective in offering helpful comments, a critical first step is to understand what makes these comments helpful. We make use of a feedback system on r/Advice that labels comments based on whether the original post author finds comments to be helpful. Based on this feedback system, we introduce a new dataset of comments, labelled with their binary helpfulness.

Helpfulness has been extensively studied based on exchanges in online support communities (Chuang and Yang, 2012; Schotanus-Dijkstra et al., 2014; Paulus and Varga, 2015; Subramani and O'Connor, 2018; McKiernan et al., 2018; Green et al., 2020). These studies found that helpfulness is associated with various characteristics such as emotional warmth, relevant knowledge, willingness to understand, empowering choice, active listening as well as sharing of similar experiences. However, these studies are solely based on qualitative interpretations and have thus far not sought to associate language features with helpfulness. To overcome this limitation, we seek to identify words that are most positively and negatively associated with helpfulness, and relate these words to characteristics of helpfulness from prior literature.

Helpfulness is closely related to empathy, as they share many characteristics such as being emotionally warm and compassionate; accepting others' frame of reference, and practising active listening (Davis, 1983; Baron-Cohen and Wheelwright, 2004; Zhou et al., 2003). We show that people's average helpfulness across all of their comments correlates with their measured empathy score. We also relate our study to literature on the language features that are associated with empathy (Sharma et al., 2020; Xiao et al., 2015; Gibson et al., 2015) and show that there is a great overlap among their

[1] https://www.reddit.com/r/Advice/

language features.

Our key contributions are:

1. We introduce and plan to openly release a novel dataset containing helpful and unhelpful comments in response to posts seeking for advice on life issues.

2. We relate helpfulness in comments that respond to posts seeking for advice on life issues to empathy.

3. We analyze the language features that are associated with helpful and unhelpful comments.

2 Related Work

Helpfulness on Online Support Communities Helpfulness has been studied in online support communities where peers can offer help and support to one another. These communities often center around a shared life situation such as chronic health conditions (Subramani and O'Connor, 2018; Green et al., 2020) and family bereavement (Schotanus-Dijkstra et al., 2014; Paulus and Varga, 2015). Several factors were emphasized in common: Peers were found more helpful when they are emotionally warm and compassionate, give others choice on a solution, willing to accept others' perspectives and experiences, practice active listening - by paraphrasing, asking questions and reflecting feelings, give pertinent advice/insights to help others to solve their problem, as well as share similar experiences (Chuang and Yang, 2012; Schotanus-Dijkstra et al., 2014; Paulus and Varga, 2015; Subramani and O'Connor, 2018; McKiernan et al., 2018; Green et al., 2020). While there has been significant work on what people find helpful, existing studies are based on qualitative themes and to the best of our knowledge, no work has been done on the language features that characterizes helpful support messages.

Language Features for Empathy Empathy is closely related to helpfulness, as many factors contributing to helpfulness (being emotionally warm and compassionate; accepting others' perspectives; practising active listening) are also associated with empathy (Davis, 1983; Baron-Cohen and Wheelwright, 2004; Zhou et al., 2003). There has been significant work on language features that characterize empathy. Sharma et al. (2020) identified that empathy is expressed in language use relating to expressing warm and compassionate emotions, communicating an understanding of others' experience, and asking more about the person's experiences. Xiao et al. (2015) and Gibson et al. (2015) found that language use relating to asking for others' perspective (*e.g.* it sounds like; do you think) are positively associated with empathy while language use that orders other around (*e.g.* you need to; please answer the) are negatively associated with empathy. Language features for empathy overlap with the features that characterize helpfulness, reinforcing the strong connection between empathy and helpfulness.

3 Dataset

Our dataset is obtained from r/Advice, which allows post authors to mark out comment(s) that they have found helpful[2]. Comments to posts with at least one empathetic comment, but were not themselves labeled as empathetic are labelled as un-empathetic. This inclusion criterion minimizes the mislabelling of comments to posts whose authors did not actively participate in labelling comments. Text from Reddit was downloaded through the Pushshift Application Programming Interface[3]. Suitable posts and all associated comments from the Advice subreddit were downloaded within 300 days (Apr 2019 - Feb 2020). Comments by the post authors and automated bots were excluded. Across the 24964 posts that were downloaded, there were 92477 associated comments (41146 empathetic). On average, each comment has 95.8 words (SD=134.5). Training/validation/test split was 80-10-10.

4 How does Helpfulness Relate to Empathy?

To determine how helpfulness relate to empathy, we calculated an aggregated metric for each user based on the proportion of their comments found to be helpful. We then correlated average user helpfulness against an established psychological measure of empathy.

Empathy Quotient Questionnaire The short form of Empathy Quotient (EQ) questionnaire (Wakabayashi et al., 2006) was used to measure empathy (details are in appendix A). Higher scores on the EQ represent higher empathy. The EQ questionnaire has high internal consistency (Cronbach's

[2]This is done using the magic word "helped", which is picked up by AdviceFlairBot

[3]https://pushshift.io/

$\alpha = 0.90$) and test-retest reliability after 12 months ($r = 0.97, p < .001$).

Participants Only users with more than 20 comments were included to minimize the likelihood that their average helpfulness was biased due to limited observations. 508 Reddit users were sent an online questionnaire through Reddit and 91 responded. Gender and age were optional to report. 86 participants reported gender (53 male and 33 female) and 83 reported age (M=33.7, SD=13.8). The mean user helpfulness is 0.5440 (SD=0.1956). Using a two-sample t-test, the distribution of EQ scores (M=24.45, SD=8.822, N=91) in this study is found to be not significantly different ($t(1850) = 0.0169, p = 0.9866$) from the sample (M=23.8, SD=8.75, N=1761) in Wakabayashi et al. (2006), demonstrating the representativeness of our sample.

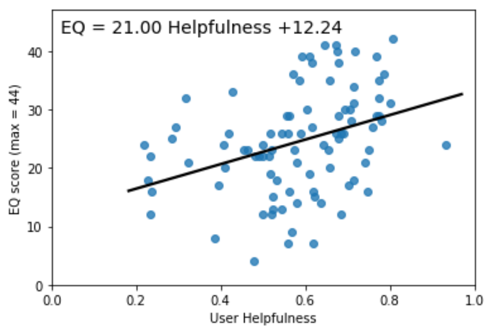

Figure 2: Empathy quotient (EQ) score against User Helpfulness

Results As illustrated in Figure 2, there is a moderate correlation effect between EQ and User helpfulness ($r(91) = 0.359, p < 0.001$). We also explored correlating User helpfulness with various subscales of the EQ, namely cognitive empathy, affective empathy and social skills based on Zhou et al. (2020). Helpfulness correlates most strongly with cognitive empathy ($r(91) = 0.355, p < 0.001$), followed by affective empathy ($r(91) = 0.261, p = 0.012$) and finally social skills ($r(91) = 0.203, p = 0.054$). This suggests that helpful commenters more often are better able to understand how the post authors think compared to how they feel or communicating it across in a social deft manner.

5 Predicting for Helpful Comments

To explore the potential for the dataset to be useful in training models to distinguish between helpful and unhelpful comments, we trained several baseline models and report their micro-average F1 scores. The performance of baseline models on this task is relatively low but similar to the performance on empathy datasets (Gibson et al., 2015; Khanpour et al., 2017; Sharma et al., 2020). The relatively low performance of baseline models on this task suggests that while recognizing helpfulness in language is trivial for typically-developing humans, they remain challenging for machines. Techniques such as commonsense reasoning (Sap et al., 2019; Bosselut et al., 2019) can be explored in the future to better capture the highly complex relationship between language and helpfulness.

	Micro-F1 (σ)
BERT	**69.2 (0.60)**
Logistic Regression	65.4 (0.55)
Naive Bayes	63.0 (0.44)
Support Vector Classifier	63.5 (0.59)
Random Forest	65.1 (0.60)

Table 1: Performance of baseline models on test set. Details of their preparation are in Appendix B

Significant Predictors of Helpfulness To understand what characterizes helpfulness in our dataset, significant predictors of helpfulness ($p < 0.05$) based on the Logistic Regression model were extracted and analysed.[4] Thematic categories that were inductively generated from these predictors are shown in Table 2 while their word clouds are available in Appendix D.

The first overarching theme is positive and friendly words. Helpfulness is positively predicted by polite, friendly-sounding and optimistic-sounding words but negatively predicted by words that indicate negative emotions. This relates to the literature findings on how uplifting and friendly online support peers are found to be more helpful. (Paulson et al., 1999; Subramani and O'Connor, 2018) Affect-related words (such as sad and tears) were previously found to be significant predictors of empathy (Gibson et al., 2015).

A second overarching theme is words relating to attempts to understand the perspective of others. Helpful commenters do so by addressing post authors directly, instead of patronizing the difficul-

[4]The dataset used to extract the most significant predictors is slightly different. Only one comment was sampled from each post and author to overcome the problem that the covariance matrix was originally non-invertible.

Direction	Themes	Words	Examples
Positive predictors	Polite, friendly sounding words	personally, friend, glad, welcome, feels, hey	Me, **personally**...I'd let it slide. He'd be That's okay I'm just **glad** that you were able to maybe text her? Be like **hey**, just wanted to say
	Optimistic sounding words	good, luck hope, hopefully yes, learned, helped forward, strong,	session with your therapist. **Good luck** **hope** something I say can help you a little! And **yes** that is dangerous and quite work that you can look **forward** to.
	Words addressing the post author directly	you	I really think **you** deserve better. **You** sound like I understand that **you** really like these guys as long as **you** feel **you** are making the best of
Negative predictors	Words indicating negative emotions	victim, kill, rid bad, depression	to be labelled as a **victim**. She might be afraid of I was internalizing every **bad** thing that happened
	Words that patronize the problem faced by the post author	dealt, wish easy, promise advice, told	it's the latter, as I **dealt** with when I was like it seems like the **easy** solution to your situation. The best **advice** I can give you though

Table 2: Thematic categories for significantly predictors of Helpfulness

ties that they face. This is also in agreement with literature on how helpfulness is associated with peers' attempt to accept others' frame of references and experiences. (Subramani and O'Connor, 2018; Green et al., 2020) Furthermore, terms indicating an inclination to find out more about the perspective of others (e.g. "do you think", "it sounds like" and "you think about") were also predictors in empathy datasets (Gibson et al., 2015; Xiao et al., 2015). Overall, the overarching themes that are predictive of helpfulness in our dataset are supported by literature on helpfulness and language features associated with empathy.

6 Human-Annotated Features for Comment Helpfulness

To better understand the capabilities and limitations of language features in capturing comment helpfulness, we manually annotated a small selection of helpful comments. For consistency of participants, we annotated 5 comments each from 91 authors who responded to our empathy quotient questionnaire. Comments were sampled using a stratified approach that results in a sampled average helpfulness to be closest possible to the author's average helpfulness score $(Pearson's\ r = 0.937, p < 0.001)$. Then we labelled each comment with one or more of the 10 possible labels based on helpfulness literature. They are 1. Highly directive, short advice 2. Dismissing concern 3. Negative terms 4. Tangential or unspecific comment 5. Share similar experience 6. Ask clarifying questions 7. Relevant knowledge 8. Emotional support 9. Recognizing difficulty 10. Tentative language. Definitions for these labels are in Appendix 4.

Using a logistic regression, we found that only the use of negative terms and tangential or unspecific comment are negatively associated with helpfulness $(p < 0.05)$ while providing relevant knowledge is positively associated $(p < 0.05)$. The use of negative terms was also captured by the logistic regression based on language use while the other two factors were not. An inspection of examples revealed that negative terms only comprises of a small set of words while those two factors require contextual semantic understanding of what is relevant knowledge to a situation and what is tangential. Future work can make use of knowledge-enhanced models (Peters et al., 2019; Clark et al., 2021) to better capture such contextual understanding.

7 Conclusion

We introduce and plan to openly release a novel dataset containing helpful and unhelpful comments in response to posts seeking for advice on life issues. Not only do we show that helpfulness of such comments is related to the commenters' empathy, we are also the first study to analyze language features that are associated with helpful and unhelpful comments. Our work can contribute towards supporting people and automated dialogue agents to offer more helpful comments to others in need.

Ethics and Broader Impact

This project has been approved by an Institutional Review Board. The use of Reddit data in this project is in alignment with the Reddit End User License Agreement and the Terms of Use for Developers. Because part of the project requires participants to respond to questionnaires, we made sure that the items were phrased sensitively so that no unintended harm would be caused. We also guided participants to make informed decisions about their participation, giving them the opportunity to withdraw any time, during and after the completion of the questionnaire. The collected information, which does not include personally identifiable information was stored securely with access restricted to the research team. We anticipate that this project can accelerate the development of models that can better detect and express helpfulness in social settings, between humans and with social dialogue agents.

References

Simon Baron-Cohen and Sally Wheelwright. 2004. The empathy quotient: An investigation of adults with asperger syndrome or high functioning autism, and normal sex differences. *Journal of Autism and Developmental Disorders*, 34(2):163–175.

Antoine Bosselut, Hannah Rashkin, Maarten Sap, Chaitanya Malaviya, Asli Celikyilmaz, and Yejin Choi. 2019. Comet: Commonsense transformers for automatic knowledge graph construction. In *ACL*.

Katherine Y Chuang and Christopher C Yang. 2012. Interaction patterns of nurturant support exchanged in online health social networking. 14(3):e54.

Thomas Clark, Costanza Conforti, Fangyu Liu, Zaiqiao Meng, Ehsan Shareghi, and Nigel Collier. 2021. Integrating transformers and knowledge graphs for Twitter stance detection. In *Proceedings of the Seventh Workshop on Noisy User-generated Text (W-NUT 2021)*, pages 304–312, Online. Association for Computational Linguistics.

Mark H. Davis. 1983. Measuring individual differences in empathy: Evidence for a multidimensional approach. *Journal of Personality and Social Psychology*, 44(1):113–126.

James Gibson, Nikolaos Malandrakis, Francisco Romero, David Atkins, and Shrikanth S. Narayanan. 2015. Predicting therapist empathy in motivational interviews using language features inspired by psycholinguistic norms. In *Proceedings of Interspeech*.

Brian M Green, Katelyn Tente Van Horn, Ketki Gupte, Megan Evans, Sara Hayes, and Amrita Bhowmick. 2020. Assessment of adaptive engagement and support model for people with chronic health conditions in online health communities: Combined content analysis. 22(7):e17338.

David M. Greenberg, Varun Warrier, Carrie Allison, and Simon Baron-Cohen. 2018. Testing the empathizing–systemizing theory of sex differences and the extreme male brain theory of autism in half a million people. *Proceedings of the National Academy of Sciences*, 115(48):12152–12157.

Y. Groen, A. B. M. Fuermaier, A. E. Den Heijer, O. Tucha, and M. Althaus. 2015. The empathy and systemizing quotient: The psychometric properties of the dutch version and a review of the cross-cultural stability. *Journal of Autism and Developmental Disorders*, 45(9):2848–2864.

Hamed Khanpour, Cornelia Caragea, and Prakhar Biyani. 2017. Identifying empathetic messages in online health communities. In *Proceedings of the Eighth International Joint Conference on Natural Language Processing (Volume 2: Short Papers)*, pages 246–251, Taipei, Taiwan. Asian Federation of Natural Language Processing.

Vladimir Kosonogov. 2014. The psychometric properties of the russian version of the empathy quotient. *Psychology in Russia: State of the Art*, 7:96–104.

Aidan McKiernan, Patrick Ryan, Eimear McMahon, Stephen Bradley, and Ellen Butler. 2018. Understanding young people's relationship breakups using the dual processing model of coping and bereavement. 23(3):192–210.

Barbara L. Paulson, Derek Truscott, and Janice Stuart. 1999. Clients' perceptions of helpful experiences in counseling. 46(3):317–324.

Trena M. Paulus and Mary Alice Varga. 2015. "please know that you are not alone with your pain": Responses to newcomer posts in an online grief support forum. 39(10):633–640.

Matthew E. Peters, Mark Neumann, Robert Logan, Roy Schwartz, Vidur Joshi, Sameer Singh, and Noah A. Smith. 2019. Knowledge enhanced contextual word representations. In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*, pages 43–54, Hong Kong, China. Association for Computational Linguistics.

Maarten Sap, Ronan Le Bras, Emily Allaway, Chandra Bhagavatula, Nicholas Lourie, Hannah Rashkin, Brendan Roof, Noah A. Smith, and Yejin Choi. 2019. Atomic: An atlas of machine commonsense for if-then reasoning. *Proceedings of the AAAI Conference on Artificial Intelligence*, 33(01):3027–3035.

Marijke Schotanus-Dijkstra, Petra Havinga, Wouter van Ballegooijen, Lynn Delfosse, Jan Mokkenstorm, and Brigitte Boon. 2014. What do the bereaved by suicide communicate in online support groups? 35(1):27–35.

Ashish Sharma, Adam Miner, David Atkins, and Tim Althoff. 2020. A computational approach to understanding empathy expressed in text-based mental health support. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 5263–5276, Online. Association for Computational Linguistics.

Sudha Subramani and Manjula O'Connor. 2018. Extracting actionable knowledge from domestic violence discourses on social media. 5(17):154807.

Akio Wakabayashi, Simon Baron-Cohen, Sally Wheelwright, Nigel Goldenfeld, Joe Delaney, Debra Fine, Richard Smith, and Leonora Weil. 2006. Development of short forms of the empathy quotient (EQ-short) and the systemizing quotient (SQ-short). *Personality and Individual Differences*, 41(5):929–940.

Bo Xiao, Zac E. Imel, Panayiotis G. Georgiou, David C. Atkins, and Shrikanth S. Narayanan. 2015. "rate my therapist": Automated detection of empathy in drug and alcohol counseling via speech and language processing. *PLOS ONE*, 10(12):1–15.

Ningning Zhou, Danni Wang, Gregory S. Chasson, Xin Xu, Jianping Wang, and Maria Izabel Lockwood. 2020. Psychometric properties of the chinese empathy and systemizing quotients in a non-clinical sample.

Qing Zhou, Carlos Valiente, and Nancy Eisenberg. 2003. Empathy and its measurement. *Positive psychological assessment: A handbook of models and measures.*, page 269–284.

A Empathy Quotient Questionnaire

Items originate from the long form of Empathy Quotient questionnaire (Baron-Cohen and Wheelwright, 2004), which is well-cited (>3500 citations) and demonstrates good validity in large (>500,000) and culturally-diverse samples (Kosonogov, 2014; Groen et al., 2015; Greenberg et al., 2018). The short form was chosen to reduce the time taken to answer the questionnaire and thereby increase the response rate. The short form is a 22-item forced-choice self-report questionnaire that can be answered on a four-point Likert Scale (Strongly Agree, Agree, Disagree, Strongly Disagree). Questions include “I often find it difficult to judge if something is rude or polite”, “I can pick up quickly if someone says one thing but means another”, and “I am good at predicting how someone will feel”. Each response can give 0, 1 or 2 points, leading to a maximum total EQ score of 44.

B Baseline Models

Each model was run with 5 different random seeds.

BERT Pre-trained BERT English-base-uncased WordPiece tokenizer was used. We fine-tuned a BERT Sequence Prediction model (English-base-uncased version with 12-layer, 768-hidden, 12-heads, 110M parameters accessed from `https://github.com/huggingface/transformers`). BertAdam optimizer was used with 0.1 epoch for warmup and learning rate of $2 * 10^{-6}$ following a search within {1,2,5} * 10^{n}, $-6 \geq n \geq -4$ using F1 as criterion. Maximum sequence length was 512 tokens, batch-size was 8 and epoch number was 2. Training took 4 hours on a Nvidia P100 GPU.

Others Text was split up into individual words and lower-cased. The number of times each word occurred in each text was then counted. Words that occurred fewer than ten times altogether were removed to minimize the effects of misspelled or rare words. Logistic Regression, Linear Support Vector Classifier, Multinomial Naive Bayes and Random Forest models were trained (accessed from `https://scikit-learn.org/stable/`) All hyperparameters were default except adjusting the number of estimators in the Random Forest model to 100. Training took negligible time (< 0.5 hours) on CPU.

C Performance of Baseline Models (Validation Set)

	Micro-F1 (σ)
BERT	**69.5 (0.52)**
Logistic Regression	65.1 (0.12)
Naive Bayes	62.9 (0.40)
Support Vector Classifier	63.5 (0.34)
Random Forest	65.2 (0.33)

Table 3: Performance of baseline models on validation set.

D Word Clouds of Significant Predictors of Helpfulness

Size of words are directly proportional to their significance of correlation.

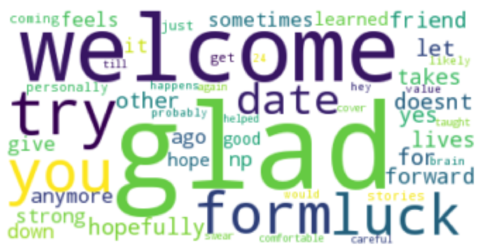

Figure 3: Significant positive predictors of helpfulness

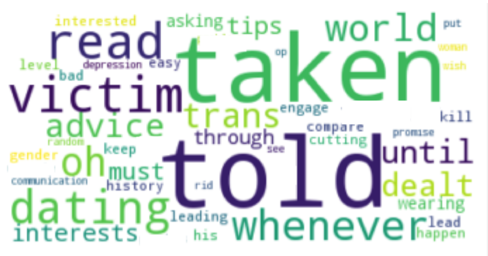

Figure 4: Significant negative predictors o helpfulness

E Labels and Descriptions for Manual Annotation of Helpfulness

Label	Description
Highly directive, short advice	Extremely short advice that are directing what the post author should do (commonly yes, no, go do this! etc)
Dismissing concern	Saying that what the post author is going through is not a big deal
Negative terms	Mentioning negative terms that the author did not bring up (crazy, psycho etc).
Tangential or unspecific comment	Mentioning random terms that has nothing to do with the author's post.
Share similar experience	Bringing up that the comment author experienced something similar as the post author
Ask clarifying questions	Asking questions to clarify what the author's situation really is. Alternatively, they can be saying "If it's situation A then . . . , otherwise if situation B then . . . "
Relevant knowledge	Bringing any knowledge to help solve the post author's specific situation (for instance, something like "you can try . . . " or "there is this resource . . . ")
Emotional support	Offering emotional comfort to the post author (something like I am sure this will get better or It's definitely not your fault)
Recognizing difficulty	Acknowledging that it's a very bad situation for the author to be in (I'm sorry that this is a really bad situation)
Tentative language	Phrasing advice as tentative suggestions – such as using "you might want to try . . . " or " I am no expert on this but . . . "

Table 4: Labels and descriptions for manual annotation of helpfulness in comments