

Paper Review #1

**Mental Health Support and its Relationship to
Linguistic Accommodation in Online Communities**

**Dept. of Computer Science & Engineering
201502755 Meeyun Kim**

Contents

1. Introduction

2. Data

3. Methods

4. Results

5. Discussion

6. Study

1. Introduction

1) About OMHCs

OMHCs(Online Mental Health Communities)

- Provide Social Support to individuals, including the **vulnerable**.
- Improve mental wellbeing like reduced **suicidal thoughts**.
- There are two types of Social Support :
Emotional Support(ES), Informational Support(IS)

1-2) About OMHCs (Social Support Types)

- **Emotional Support(ES), Informational Support(IS)**

empathetic messages → **direct ES**

exposed to others having similar experiences → **indirect ES**

receiving helpful information, advice related to treatment, medication /
building social capital → **IS**

- In this paper, researchers **automatically** detect levels of **ES** and **IS** in posts.

2) About Linguistic Accommodation

- Linguistic Accommodation Theory concerns
 - (1) the **behavioral changes** that people make **to attune** their communication to their partner,
 - (2) the extent to which people perceive **their partner as appropriately attuning to them.**

(Wikipedia, Communication accommodation theory)

- "When people interact they adjust their speech, their vocal patterns and their gestures, to accommodate with others."

(Howard Giles, Speech Accommodation theory)

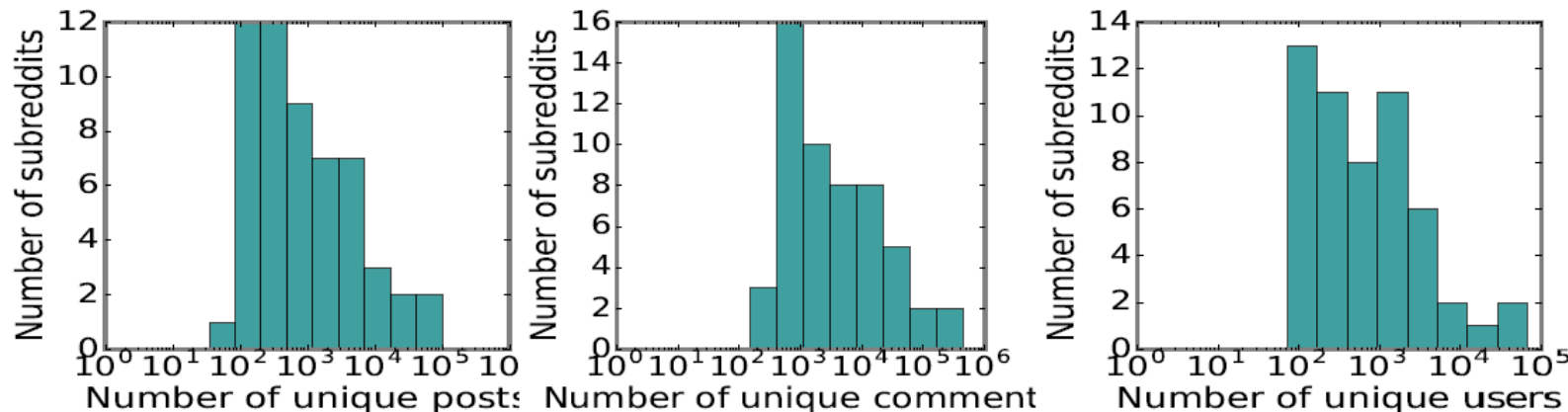
3) Relationship between OMHCs' Support and Linguistic Accommodation

- **Main topic of this paper** : How do members' linguistic accommodation practices impact the support provisions in OMHCs?
- The identity of OMHCs is defined by **norms**.
- Across diverse OMHCs, researchers found that **with an increase in the linguistic accommodation in a post, there is an increase in the ES or IS.**

2. Data

1) Data Acquisition

- Researcher used accessible data from **55 Reddit OMHCs**.
→ to compile a list of OMHCs, researchers employed **Snowball sampling**.
(keyword set : support, counseling, mental health, mental, trauma, abuse, depression, suicide, therapy, coping)
- **Reddit data archive** (can use in Google's BigQuery)
→ Archive provided **358,409 posts** and **1,832,702 comments**.
- The average **number of subscribers** is **132 to 189,922**.



2) Categories of OMHCs

- Researchers **categorized all OMHCs** based on the broader topics and mental health issues they focus on.
- Used a **two-step unsupervised learning based machine-human framework.** (human labeling + clustering)

2) Categories of OMHCs

- ① **Clustering** → using **k-means clustering algorithm** to perform initial clustering on the **n-grams(n=3)** of the posts shared in the 55 OMHCs.
- ② **Human labeling** → two annotators independently refined these machine-labeled clusters and assigned them label. **(with semi-open coding approach)**

Annotators borrowed the conceptualizations of different mental health conditions in **DSM-5**.

2) Categories of OMHCs

Community Category	Communities	#Posts	#Comments	#Users
<u>Trauma & Abuse (C₁)</u>	<i>r/abuse, r/adultsurvivors, r/afterthesilence, r/Anger, r/bullying, r/CPTSD, r/domesticviolence, r/emotionalabuse, r/ptsd, r/PTSDCombat, r/rapecounseling, r/StopSelfHarm, r/survivorsofabuse, r/SurvivorsUnited, r/traumatoolbox</i>	11,213	44,486	8,661
<u>Psychosis & Anxiety (C₂)</u>	<i>r/Agoraphobia, r/Anxiety, r/BipolarReddit, r/BipolarSOs, r/BPD, r/dpdr, r/psychoticreddit, r/MaladaptiveDreaming, r/Psychosis, r/PanicParty, r/schizophrenia, r/socialanxiety</i>	70,696	374,072	41,580
<u>Compulsive Disorders (C₃)</u>	<i>r/calmhands, r/CompulsiveSkinPicking, r/OCD, r/Trichsters</i>	8,032	35,948	5,340
<u>Coping & Therapy (C₄)</u>	<i>r/7CupsofTea, r/BackOnYourFeet, r/Existential_crisis, r/getting_over_it, r/GriefSupport, r/helpmecope, r/hardshipmates, r/HereToHelp, r/itgetsbetter, r/LostALovedOne, r/offmychest, r/MMFB, r/Miscarriage, r/reasonstolive, r/SuicideBereavement, r/therapy</i>	100,248	510,168	77,031
<u>Mood Disorders (C₅)</u>	<i>r/depression, r/depressed, r/ForeverAlone, r/GFD, r/lonely, r/mentalhealth, r/Radical_Mental_Health, r/SuicideWatch</i>	168,220	868,028	112,915

- Annotators finally identified five categories of mental health communities: **Trauma & Abuse, Psychosis & Anxiety, Compulsive Disorders, Coping & Therapy, Mood Disorders.**

3. Methods

1) Support Classification

- Researchers infer **the levels of ES and IS** from the comments made on the posts in OMHC dataset. **(Two-class characterization)**
- Two annotators separately coded each of them for the level of ES and IS, using a **three-point Likert Scale**. (1 = least supportive, 3 = most supportive)
- The **inter-rater agreement metric Cohen's kappa** was found to be **0.879** and **0.876** for IS and ES respectively.
- In coded sample of 400 posts, there were **185, 131, and 81 posts** with **IS** scores 1, 2, and 3, **135, 134, and 128 posts** with **ES** scores 1, 2, and 3 respectively.

1) Support Classification

- Using these annotated comments as training data, researchers built two **tertiary classification models**(support score classes : 1, 2, 3)
- Researchers employed **LIWC** to calculate classification features in comment text that could be predictive of ES or IS.
- IS classification model included an **extra binary indicator** feature corresponding to the presence of **URLs**.

1) Support Classification

- **Multinomial logistic regression model**

→ Best performance for predicting the **ES** of a comment(accuracy, F-1 score)

- **Support Vector Machine(SVM) classifier**

→ Best for predicting **IS**.

- Both models was evaluated using **k-fold cross-validation(k=10)**

- **ES classifier's** mean accuracy was **59%** (better than baseline accuracy of 35%),
IS classifier's mean accuracy was **72%** (better than baseline accuracy of 60%)

2) Modeling the Link between Support and Accommodation

Linguistic Style Matching(LSM)

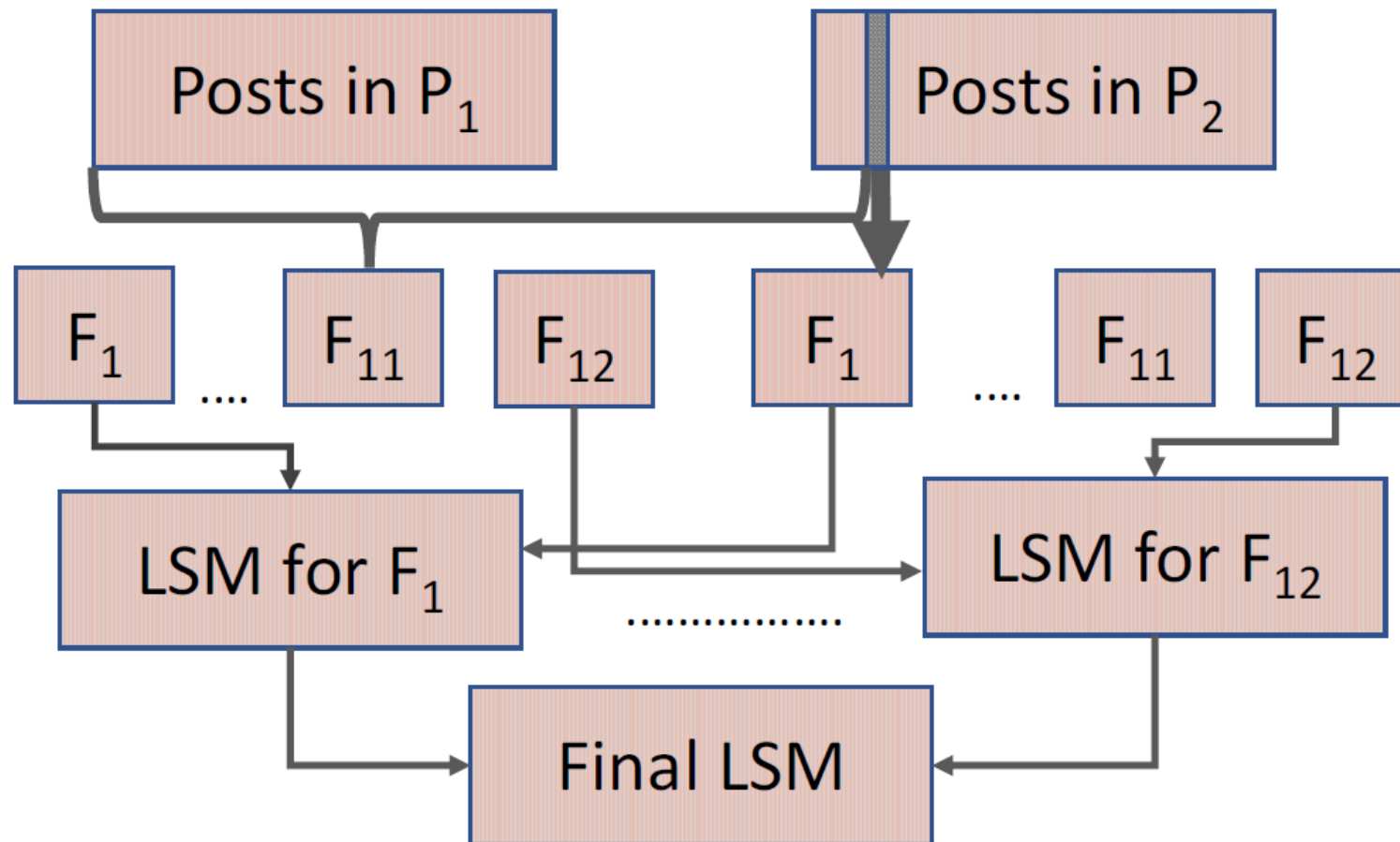
- Researchers present the method for measuring linguistic accommodation with **Linguistic Style Matching(LSM)**.
- LSM considers the rate of use of **function words** in an individual's speech. It help identify relationships between language and social psychological states.

2) Modeling the Link between Support and Accommodation

- ① Linguistic accommodation is established **over time**.
 - Researchers split posts in each community into two parts, P_1 and P_2 , based on the median of their **timestamps**.
 P_1 has a smaller timestamp value than P_2 .
- ② Then, Researchers began by calculating the proportion of function words in P_1 and P_2 posts using the 12 categories in LIWC.

2) Modeling the Link between Support and Accommodation

③



3) Statistical Models

- Researchers developed two sets of **nested multinomial logistic regression** models, one for ES and the other for IS.
- **Response variable** : **support score** of a post(1, 2, and 3)
- **Set of models** :
 - ① **Control model** – quantifies the relationship between confounding predictor variables and the support scores.
 - ② **LSM model** – includes the **LSM score** of a post as an **additional predictor variable**.

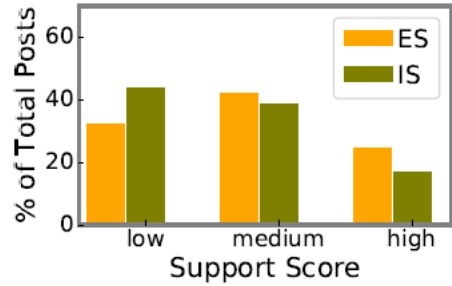
3) Statistical Models

Confounding predictor variables

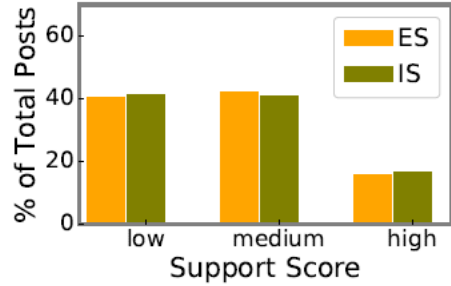
- ① **Content** : a post's topical content. (used LIWC dictionary)
- ② **Post Length** : lengthy posts receive greater support.
(white spaced words used as a control variable)
- ③ **Self-Disclosure**
- ④ **Author Tenure and Familiarity** :
 - determine the **tenure** of post's author by timestamp.
 - quantify a post author's **familiarity** by number of posts made by them.
- ⑤ **Throwaway Account** : **anonymity** of author impacts level of support.
- ⑥ **Size of the Community** : number of subscribers used as a variable.

4. Results

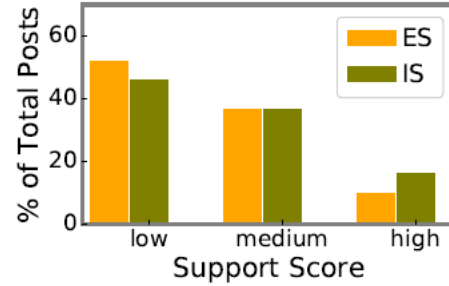
1) Support Classification



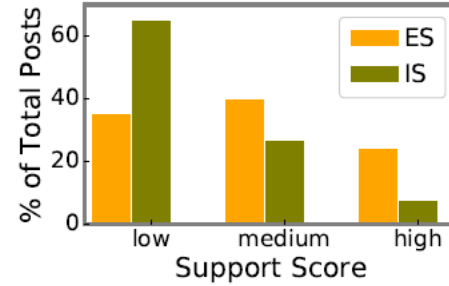
(a) Trauma & Abuse



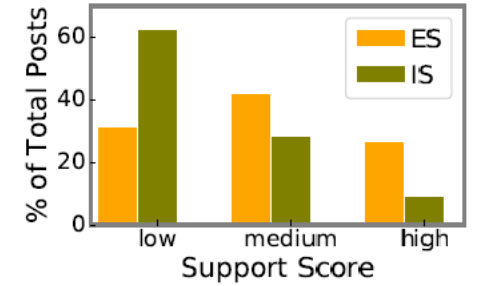
(b) Psychosis & Anxiety



(c) Compulsive Disorders



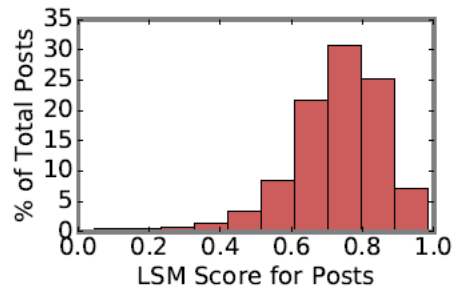
(d) Coping & Therapy



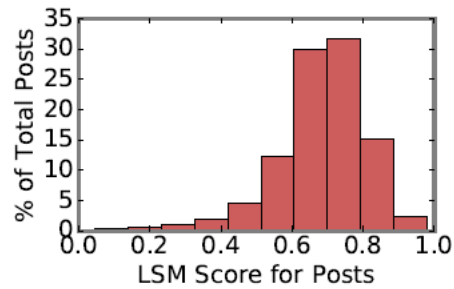
(e) Mood Disorders

- **Trauma & Abuse, Coping & Therapy, Mood Disorders**
→ **ES** measured higher.
- **Compulsive Disorders**
→ **IS** measured higher.
- **Psychosis & Anxiety**
→ **ES** and **IS** measured similarly.
- This could be due to the **needs of individuals** approaching these communities or **the goal of the communities** itself.

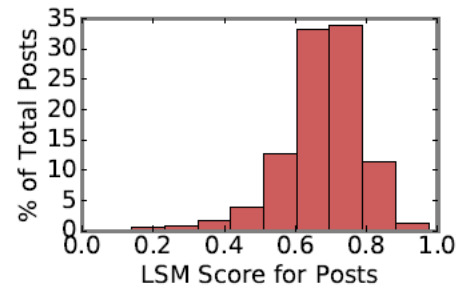
2) Examining Support and LSM



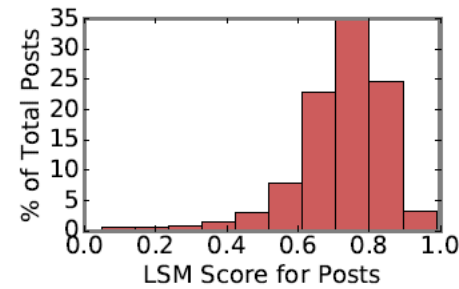
(a) Trauma & Abuse



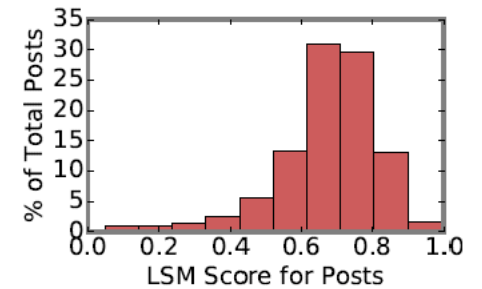
(b) Psychosis & Anxiety



(c) Compulsive Disorders



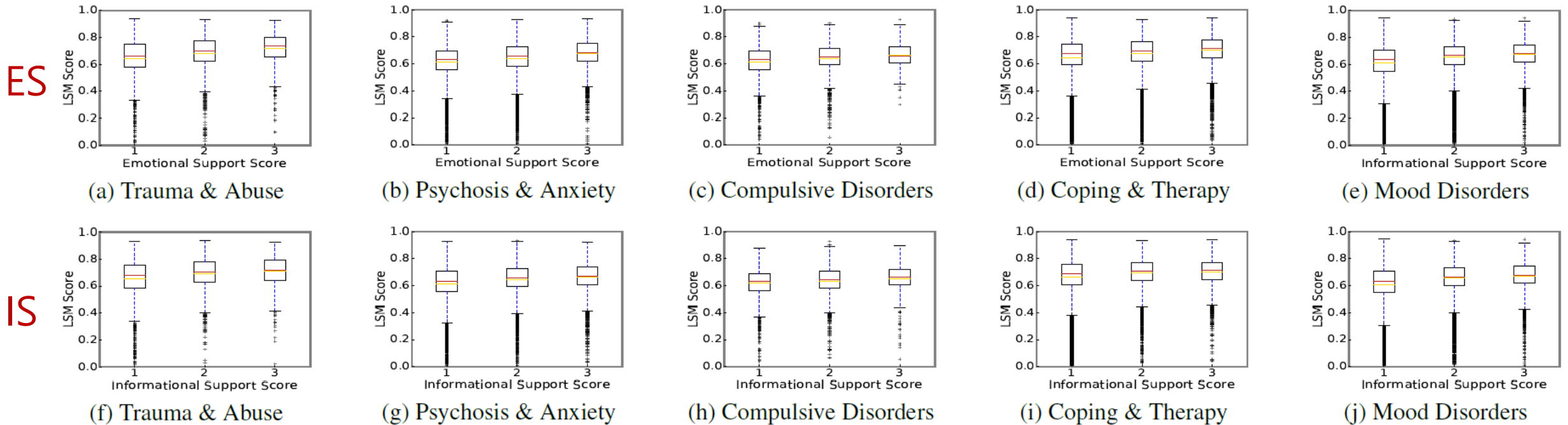
(d) Coping & Therapy



(e) Mood Disorders

- Researchers found that the mean LSM for all community categories lies in the range of 0.63 to 0.68.
- This indicates that people tend to **conform with the language conventions** of a community and **mimic the style of writing** of the rest of the community.

2) Examining Support and LSM



- As shown in the **box-and whisker plot**, the increase in both IS and ES in any community category, there is a consistent increase in linguistic accommodation.
- But the increase of **ES in Trauma & Abuse, Coping & Therapy** and increase of **IS in Trauma & Abuse, Mood Disorders** are noticeable.

2) Examining Support and LSM

- Researchers divide the posts into **two groups** : high support and high LSM, low support and low LSM.
- Researchers used **SAGE** which is a generative model of text with class label.
- Results from SAGE Analysis show frequent and infrequent words used in both group (high support and high LSM / low support and low LSM)

Informational Support (IS)	
High LSM, High Support	
C ₁	depressive, music, insight, episode, classes
C ₂	pains, blood, nausea, <u>xanax</u> , prozac
C ₃	<u>insulin</u> , stab, <u>luvox</u> , <u>abnormal</u> , scalp
C ₄	medication, motivation, doctor, treatment, insurance
C ₅	psychiatrist, symptoms, medication, prescribed, disorder
Low LSM, Low Support	
C ₁	black, scars, name, write, girls
C ₂	victories, sharing, app, dpdr, message
C ₃	simply, gift, sides, learn, pull
C ₄	moon, fuck, you, your, shame
C ₅	bills, women, loans, birthday, girls
Emotional Support (ES)	
High LSM, High Support	
C ₁	awkward, disgusted, wonderful, trusted, hated,
C ₂	relationship, distance, dating, failure, hurts, deserve
C ₃	extremely, ruminations, fruit, plagiarism, curiosity
C ₄	suicidal, blame, dead, feelings, childhood
C ₅	together, killing, hurting, husband, loved
Low LSM, Low Support	
C ₁	treatments, ice, plan, book, bear,
C ₂	coffee, caffeine, disease, diet, asleep,
C ₃	math, amp, contamination, posted, driving
C ₄	moon, fuck, dollars, water, teeth
C ₅	loans, steam, bills, information, curious

3) Nested Multinomial Logistic Regression

Goodness of fit

- **Control, LSM model** showed improvement in variance compared to null model.
- **LSM model** approximately follows a **chi-squared distribution**.
- Mood Disorder(C_5) communities LSM is highly predictive of IS.

Community category (# obs.)		Null	Control	LSM
<i>Informational Support (IS)</i>				
	df	0	88	90
	deviance	11543	11097.4	11044.8
C_1 (5,606)	χ^2		445.6***	52.6***
	deviance	72552	69186	68880
C_2 (35,348)	χ^2		3366***	306***
	deviance	8207.6	7811	7799.4
C_3 (4,016)	χ^2		396.6***	11.6***
	deviance	83290	79622	79358
C_4 (50,124)	χ^2		3668***	264***
	deviance	146544	140386	139672
C_5 (84,110)	χ^2		6158***	714***
<i>Emotional Support (ES)</i>				
	df	0	88	90
	deviance	12043.8	11457.2	11415.8
C_1 (5,606)	χ^2		586.6***	41.4***
	deviance	72098	68456	68334
C_2 (35,348)	χ^2		3642***	122***
	deviance	7556.6	7260.8	7242.8
C_3 (4,016)	χ^2		295.8***	18***
	deviance	108102	104486	104218
C_4 (50,124)	χ^2		3616***	268***
	deviance	181788	175434	175028
C_5 (84,110)	χ^2		6354***	406***

3) Nested Multinomial Logistic Regression

	Informational Support (IS)				Emotional Support (ES)			
	β Weights		Odds Ratios		β Weights		Odds Ratios	
	$y = 2$	$y = 3$	$y = 2$	$y = 3$	$y = 2$	$y = 3$	$y = 2$	$y = 3$
C_1	1.75	2.98	<u>5.74</u>	<u>19.72</u>	1.29	2.77	<u>3.64</u>	<u>16.01</u>
C_2	1.642	2.85	5.15	17.33	0.83	1.99	2.28	7.29
C_3	0.71	1.82	<u>2.04</u>	<u>6.18</u>	0.89	2.79	<u>2.45</u>	<u>16.35</u>
C_4	1.68	2.17	5.39	8.75	0.93	2.26	2.52	9.53
C_5	1.80	2.92	<u>6.07</u>	<u>18.48</u>	0.73	2.01	<u>2.08</u>	<u>7.43</u>

(Low support score(=1) is the referent group for both IS and ES regression models.)

- Odds ratio greater than 1 indicates that the risk of the support score being 2 or 3 relative to the risk of the support score being 1 increases as LSM increases.
- It means that the chances of **getting more support with the increase in LSM score is higher in case of IS than in ES.**

5. Discussion

1) Theoretical Implications

Support Matching and Provisioning

- Certain communities direct more ES, while some others provide more IS.
 - Coping with a mental illness like **depression** may need **more ES**.
 - The posters in the **Compulsive Disorders** communities tend to **seek advice around their condition**, so **IS** is more effective.
- This indicates that in spite of all these communities being related to mental health **they cater to different social support goals**.

1) Theoretical Implications

Conformance and Support

- The most notable finding of this work is that all types of **social support in OMHCs is positively associated with linguistic accommodation.**
- Linguistic alignment is required to receive a good level of support in OMHCs.

1) Theoretical Implications

The Tension Between Accommodation and Support

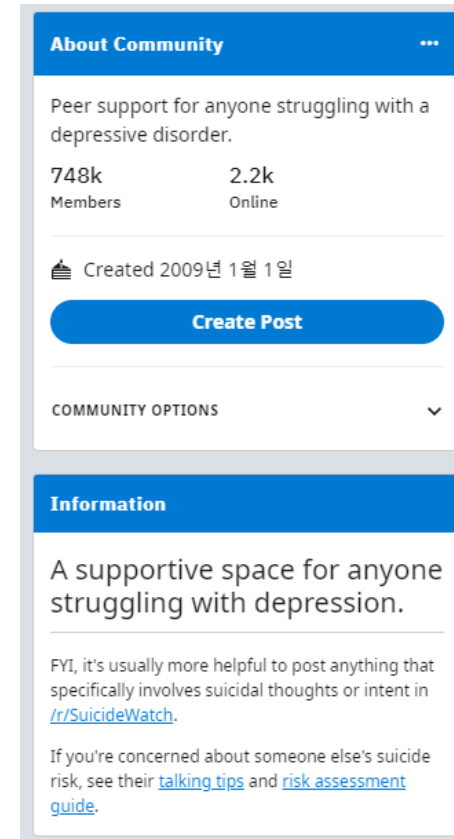
- **Should linguistic alignment be an absolute necessity to receive help**, even if somebody expresses critical needs, such as risk of self-injury?
 - Should **less conforming posters not receive sufficient support** even if they genuinely **require immediate attention**?
- There can be two answers to these questions :
- ① allowing non-conforming support seekers gain explicit awareness of the underlying norms. **(for support seekers)**
 - ② equipping support providers with better tools to easily direct help to individuals with critical psychological needs. **(for support provides)**

2) Design Implications

Influencing Community Design

- The sidebar pane of a subreddit's landing page is often used in OMHCs to inform members about the purpose of that community along with the type of content that is permissible to be shared in that community.

- However, these do not provide a simple way for support seekers, often already challenged by critical situations.



/r/depression rules

1. Text Posts only.

- Articles (including personal blogs) and media must be linked in a text submission and accompanied by a full account of your personal experience with them and how they affected your depression.
- If your resource is of primarily scientific, philosophical or sociopolitical interest, please post it elsewhere, e.g. [/r/psychology](#), [/r/science](#), [/r/philosophy](#), [/r/mentalhealth](#).
- This also means no fundraising or activism, please.
- Any surveys or research related posts *must* be pre-approved by the mod team. Approved research posts will have a tag that only moderators can apply.
- Do not debate or sensationalise current events or hot-button topics because they happen to have a "depression angle".

2. Do not post or respond with:

- "Tough love" of any description.
- General uplifting or "it gets better" messages. **Encouragement is *not* helpful unless it integrates real, personal understanding of the OP's feelings and situation.**
- Anything explicit or inciting related to suicide or self harm.
- Claims about the efficacy of any treatment or self-help strategy *including religion*.

3. Please **don't diagnose others or advocate for or against specific treatments or self-help strategies.**

- In response to someone who's asked for help, it's okay to **describe** how something worked (or didn't) for you, but it's absolutely not okay to **prescribe** for anybody else.
- Our focus is on **support rather than "fixes"**. If you are looking for a more solution-oriented community, check out [/r/getting_over_it/](#) or [/r/GetMotivated](#).

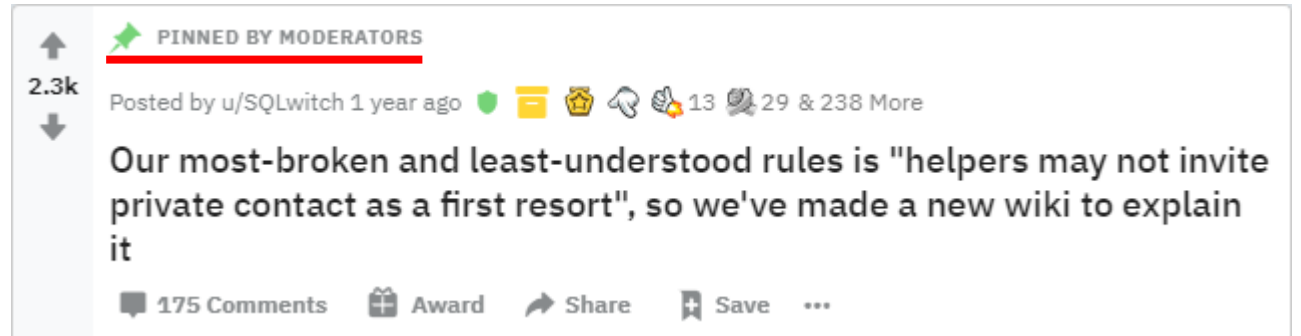
4. **Don't say that you're here to help, show it**

- by responding supportively to our OPs.
- Empathy, support, and feedback are usually more helpful than advice.

2) Design Implications

Influencing Community Design (Cont'd)

- Researchers proposed alternate design which can promote **good example posts** at the top of the landing page of communities.



- Good example posts mean historical posts with both high linguistic accommodation and high ES or IS.

- For Reddit, this can be implemented via **sticky comments**, a post with a pinned mark.

2) Design Implications

Assisting Support Providers

- In communities such as like r/SuicideWatch, OMHCs need to include design capabilities that can '**override**' the implicit or explicit need for greater linguistic accommodation in receiving adequate social support.
- Researchers suggest the development of intervention tools as a solution, that can issue **timely alerts to active support providers**.
- The support providers can then reach out to posts that need help(**low linguistic accommodation and which have received little support posts**).

3) Limitations and Future Work

- Support classifier which proposed by researchers showed satisfactory performance and agreed well with expert annotations.
- However, **there could also be multiple factors**, beyond linguistic accommodation that could impact whether a post shared in an OMHC receives adequate support.
- Although researchers found a positive link between linguistic accommodation and support, **they cannot be certain if this is causal in nature.**

6. Study

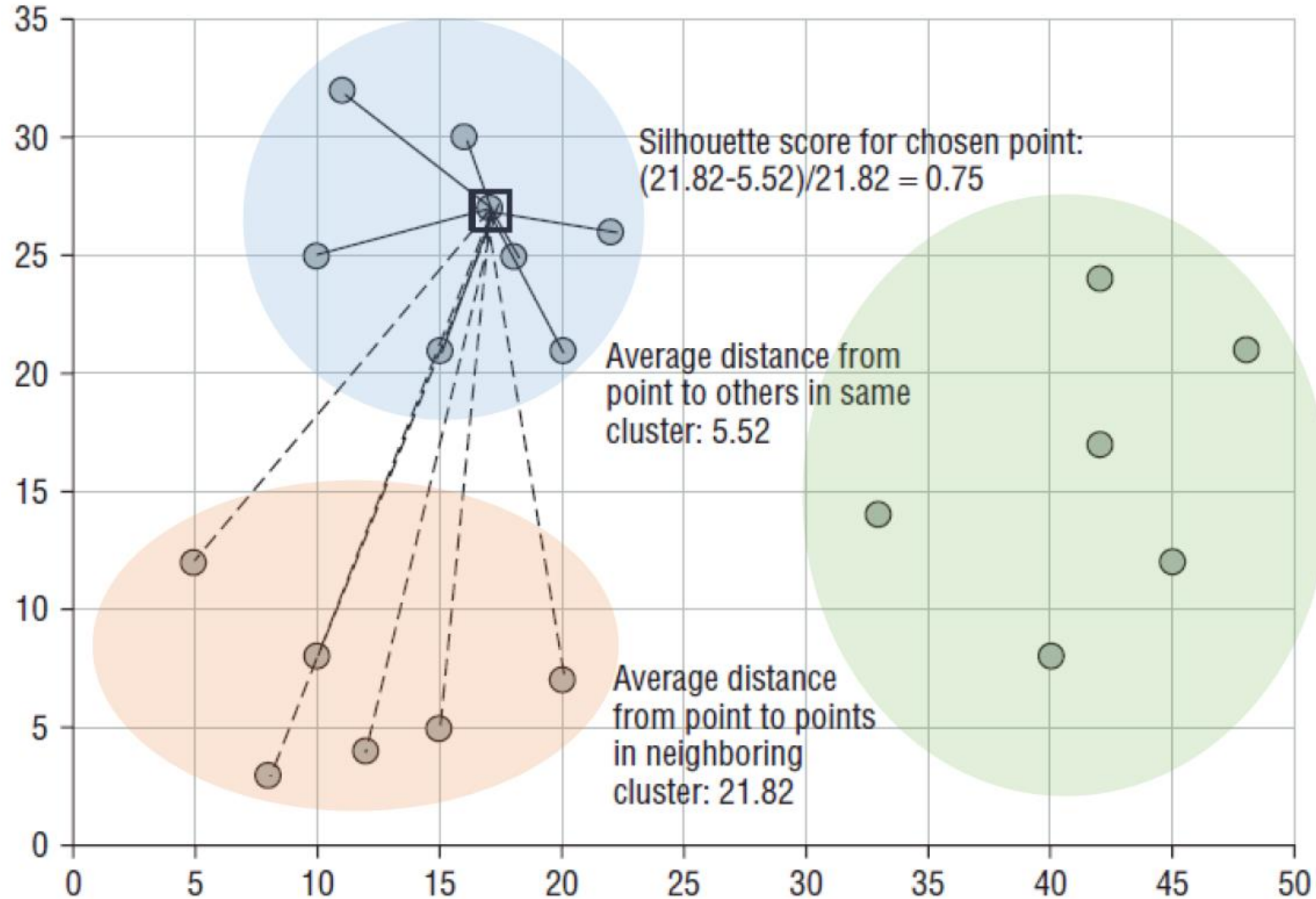
1) Silhouette Coefficient

- a method of interpretation and validation of consistency within clusters of data.
- The silhouette value is a measure of **cohesion** and **separation**.
- $a(i)$ (cohesion) :
- $b(i)$ (separation) :

$$a(i) = \frac{1}{|C_i| - 1} \sum_{j \in C_i, i \neq j} d(i, j)$$

$$b(i) = \min_{k \neq i} \frac{1}{|C_k|} \sum_{j \in C_k} d(i, j)$$

1) Silhouette Clustering



1) Silhouette Clustering

- Define a silhouette (value) of one data point i

$$s(i) = \frac{b(i) - a(i)}{\max\{a(i), b(i)\}}, \text{ if } |C_i| > 1$$

$$-1 \leq s(i) \leq 1$$

and

$$s(i) = 0, \text{ if } |C_i| = 1$$