The Relationship of User-Generated Content with Responsiveness and Commitment in Online Communities

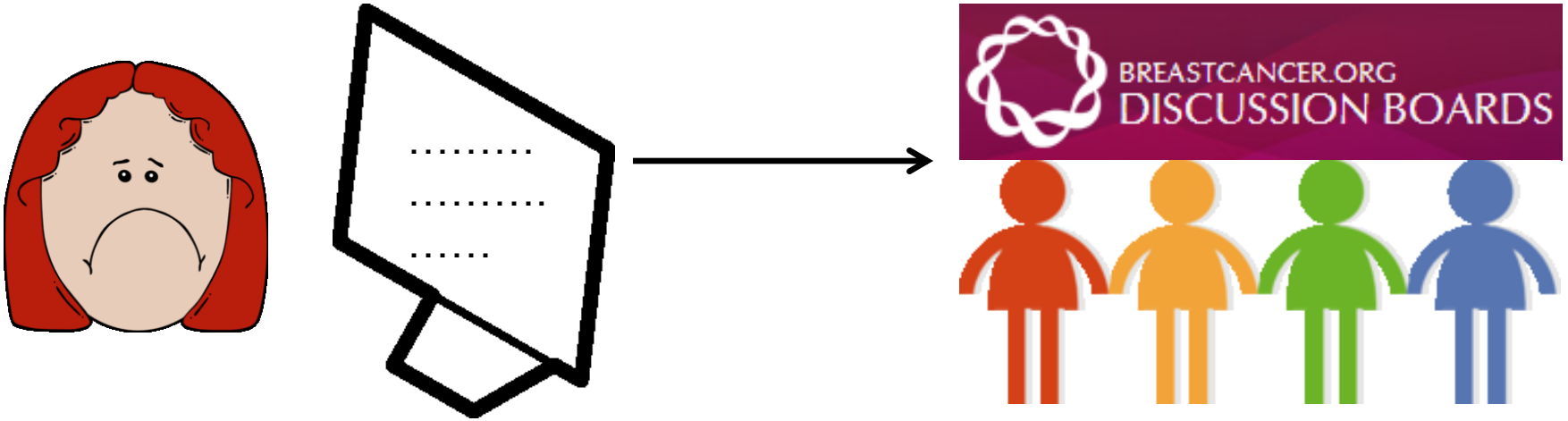
Yi-Chia Wang

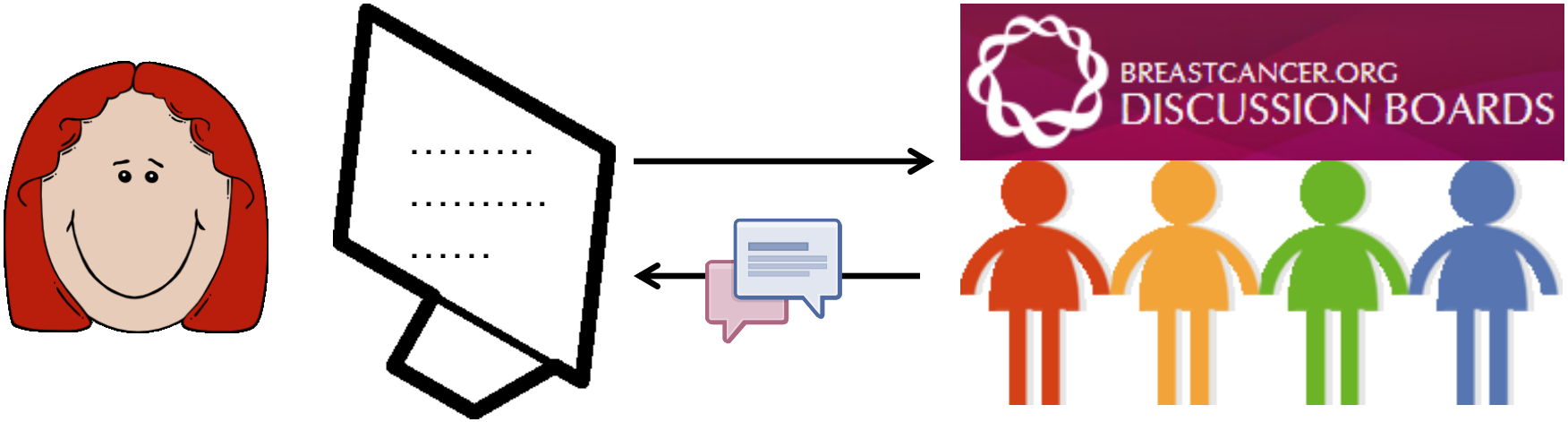
Language Technologies Institute

Carnegie Mellon University

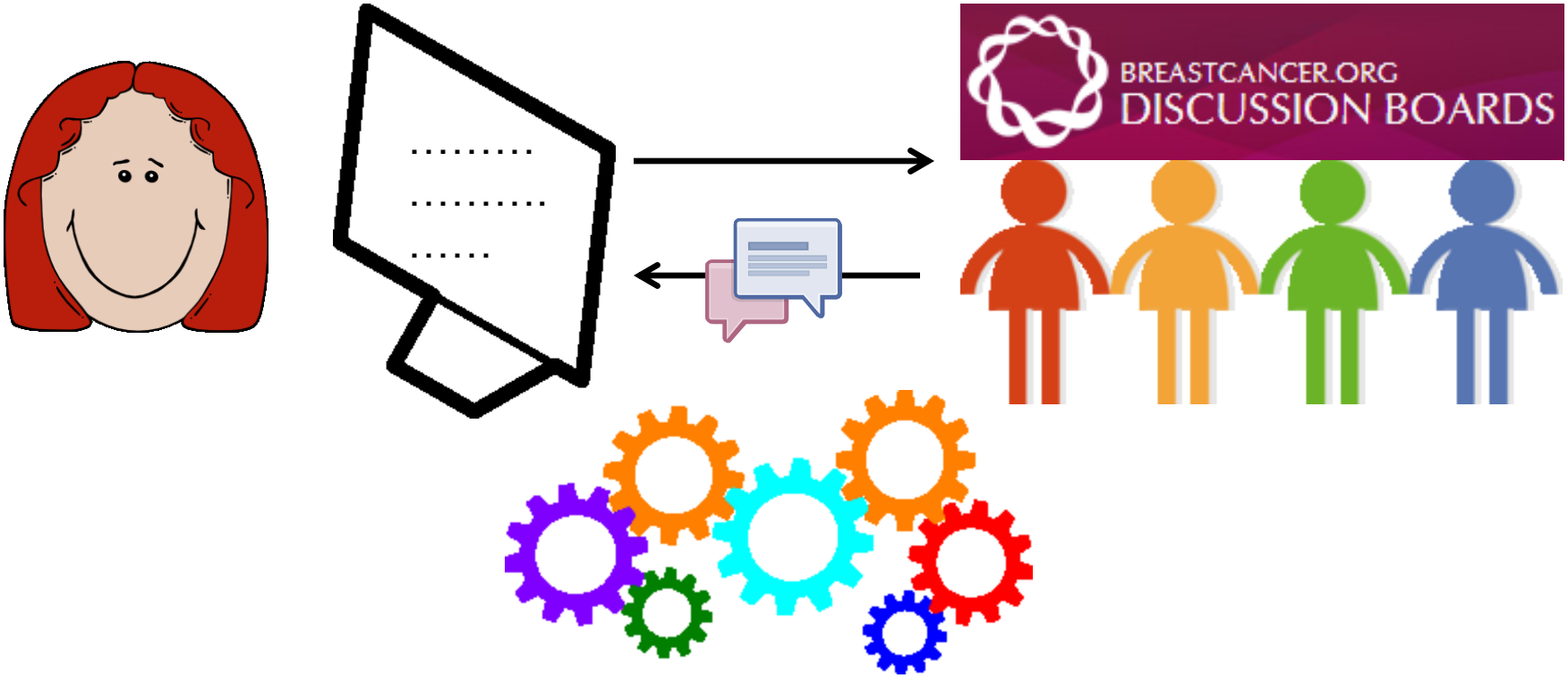
Advisor: Robert (Bob) Kraut, *HCII*


I don't even know where to begin
I'm so upset and very scared
What is Chemotherapy?
Where to find support and someone to talk to?

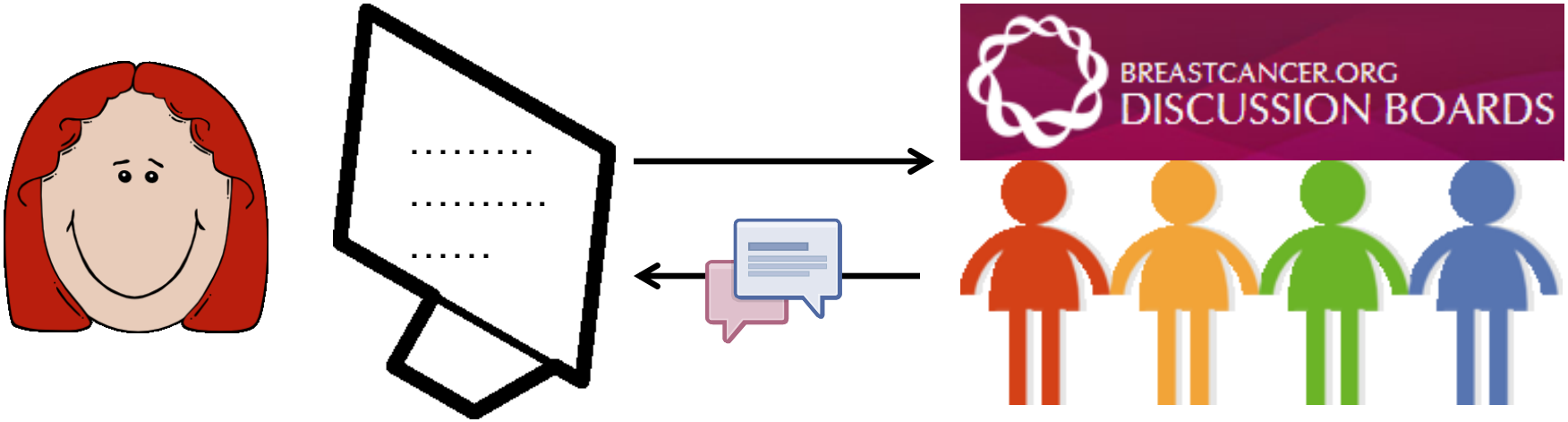

BREASTCANCER.ORG
DISCUSSION BOARDS

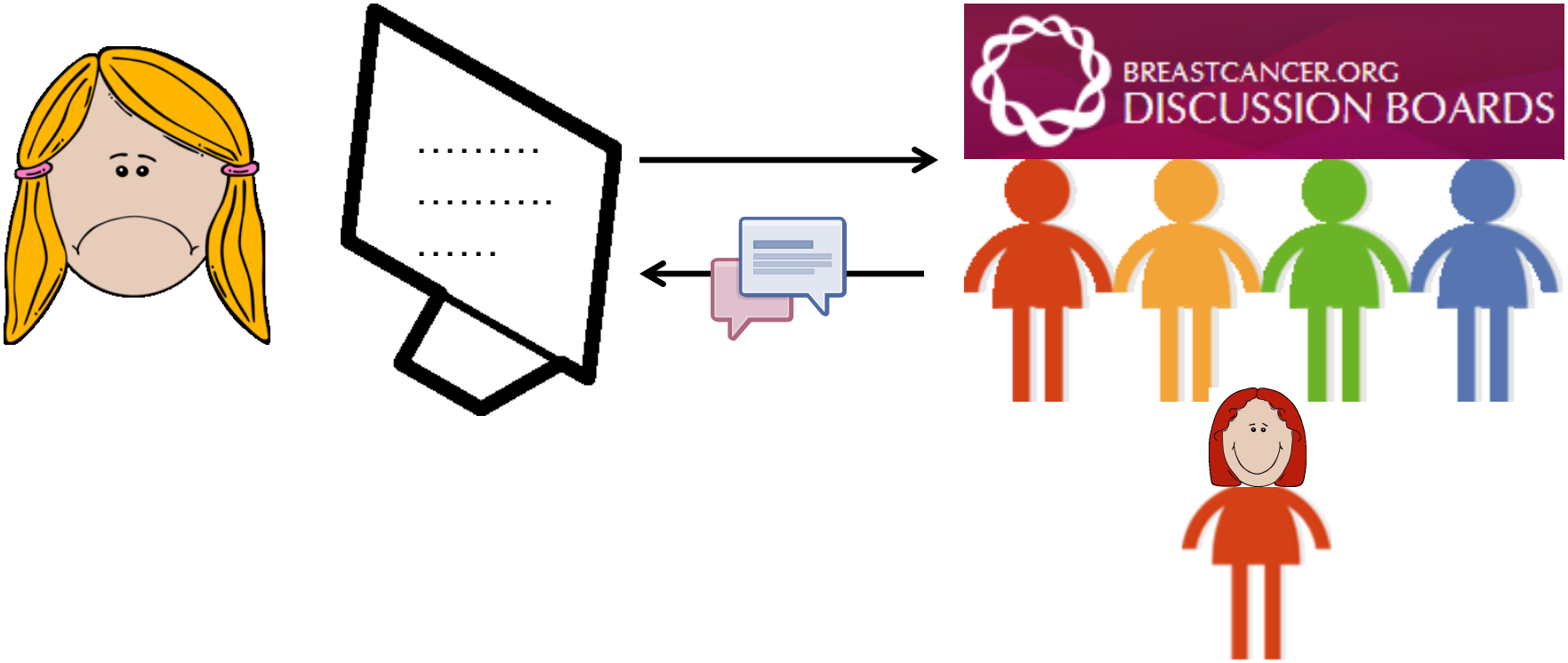


Q1: How to elicit responses that satisfy one's need in online groups?



- Machine learning approach (Bottom-up)
- Social science approach (Top-down)


BREASTCANCER.ORG
DISCUSSION BOARDS

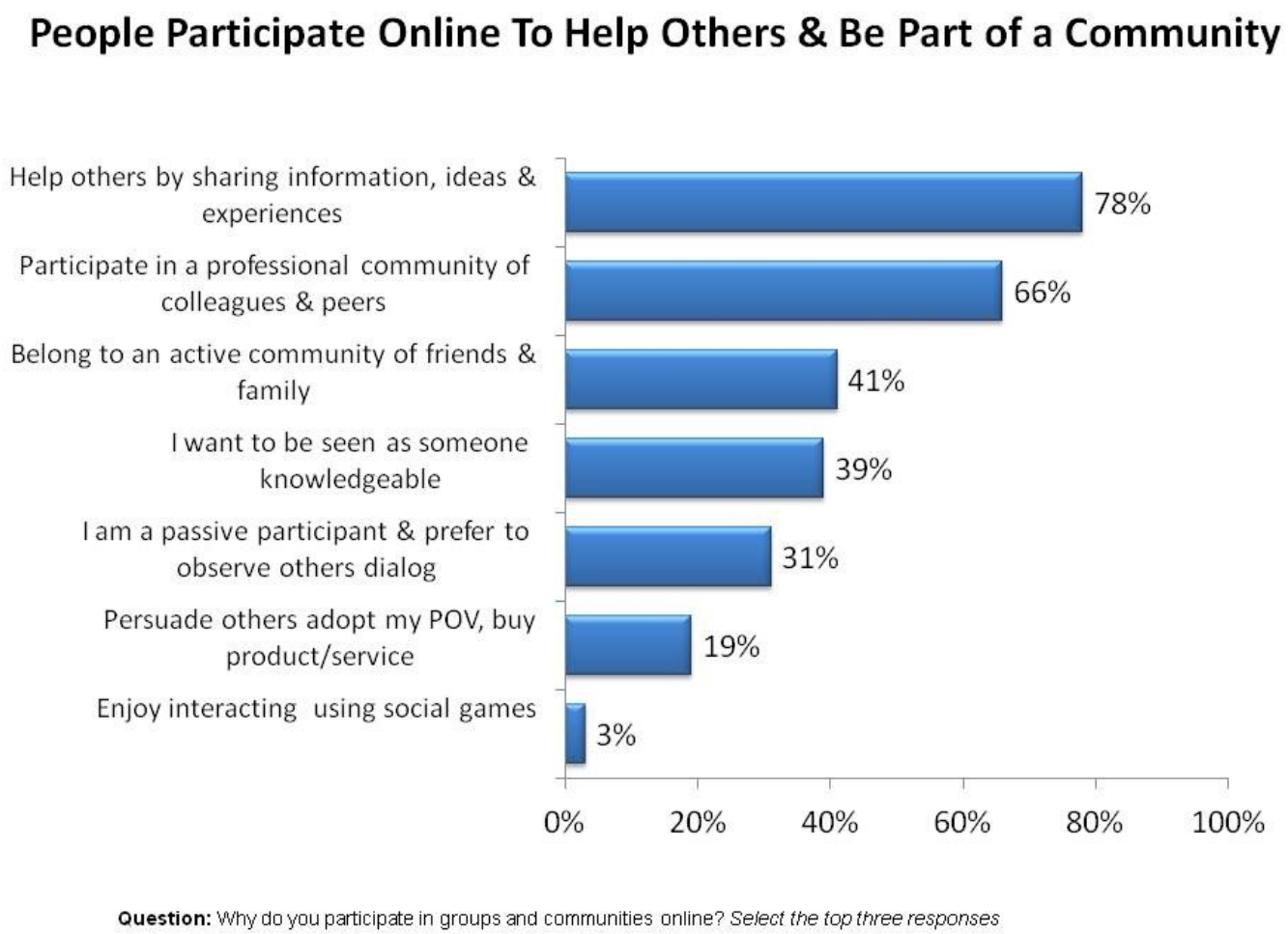


Q2: What are the mechanisms that affect one's participation in online groups?

Success of Online Communities

- Responsiveness: the community's willingness to respond
  - In Usenet ~ 40% of initial posts get no response
- Commitment: members' willingness to come back to the community
  - 70% of newcomers who initially post on Usenet never post again

Why Do People Hangout Online?



http://socialmediatoday.com/vanessa-dimauro/646626/nearly-80-people-participate-online-community-help-others

Responsiveness

JOURNAL OF MEDICAL INTERNET RESEARCH

How to Get Support?

Understanding the Conversation Dynamics of Social Support

Yi-Chia Wang[1], Robert Kraut[2], John M. Levine[3]

[1]Language Technologies Institute, Carnegie Mellon University, Pittsburgh, USA

[2]Human-Computer Interaction Institute, Carnegie Mellon University, Pittsburgh, USA

[3]Department of Psychology, University of Pittsburgh, Pittsburgh, USA

ABSTRACT

Background: A high percentage of people with serious diseases participate in online health support groups with the goal of obtaining help. However, little systematic research has been devoted to investigating how participants in these groups obtain the help they seek. Furthermore, most studies on communication in support groups are based on hand-coding relatively small samples of messages.

Objective: This paper had three goals. The first was to propose a model of the dynamic process of social support communication in online health support groups. The second was to develop a methodology using automated computer models and machine learning algorithms to analyze conversations in online support groups. And the third was to test the adequacy of our model using machine-coded and hand-coded data.

Methods: We used Amazon Mechanical Turk workers to construct a hand-coded dataset indicating the extent to which a message is self-disclosing, asks questions, seeks emotional and informational support, and provides emotional and informational support. Using human-coded data as ground truth, we applied natural language processing techniques and machine learning algorithms to build models for detecting self-disclosure, questions, support elicitation, and support provision in online conversations.

Results (make sure to include relevant statistics here, such as sample sizes, response rates, P-values or Confidence Intervals. Do not just say "there wer

The analysis on the human coded data shows and negative sel
disclosures (0.38***) are effective ways for ge indicates that th

Commitment

To Stay or Leave? The Relationship of Emotional and Informational Support to Commitment in Online Health Support Groups

Yi-Chia Wang
Language Technologies Institute
Carnegie Mellon University
Pittsburgh PA 15213
yichiaw@cs.cmu.edu

Robert Kraut
Human-Computer Interaction Institute
Carnegie Mellon University
Pittsburgh PA 15213
robert.kraut@cs.cmu.edu

John M. Levine
Psychology
University of Pittsburgh
Pittsburgh PA 15260
jml@pitt.edu

ABSTRACT
Today many people with serious diseases use online support groups to seek social support. For these groups to be sustained and effective, member retention and commitment is important. Our study examined how different types and amounts of social support in an online cancer support group are associated with participants' length of membership. We first built machine learning models to automatically identify the extent to which messages contained emotional and informational support. Agreement with human judges was high ($r > 0.76$). We then used these models to measure the support exchanged in 1.5 million messages. Finally, we applied quantitative event history analysis to assess how exposure to emotional and informational support predicted group members' length of subsequent participation. The results demonstrated that the more emotional support members were exposed to, the lower the risk of dropout. In contrast, informational support did not have the same strong effects on commitment. We speculate that emotional support enhanced members' relationships with one another or the group as a whole, whereas informational support satisfied members' short-term information needs.

Author Keywords
Commitment; Online communities; Social support; Natural language analysis; Applied machine learning.

ACM Classification Keywords
H5.3. Information Interfaces and Presentation: Group and Organization Interfaces: Asynchronous interaction, Computer-supported cooperative work, Evaluation/methodology, Web-based interaction.

INTRODUCTION
A large number of American Internet users participate in online health support groups to obtain informational and/or emotional support [8, 9]. A large fraction of these online support groups deal with cancer [11].

Although online support groups are popular, the scientific jury is out regarding their effectiveness in helping participants deal with health problems [20]. It is highly likely that the effectiveness of such groups depends on the communications that members exchange with one another, but surprisingly little systematic research has been devoted to specifying how the quality and quantity of such communications affect groups' outcomes and members' health-related outcomes (see [16, 25, 28] for exceptions).

This paper focuses on member retention and commitment, which are important to individual members and to the maintenance and success of the group as a whole. People who stay in an online support group longer are more likely to receive whatever benefits it provides. Moreover, members are resources in online groups. They share information, provide help, and form social ties with others. Over time, they shift from receiving support to providing it to others [27]. Continued participation, however, cannot be taken for granted. Evidence obtained in a wide variety of online groups indicates that a substantial number of participants drop out before they could plausibly contribute any benefits to the groups or receive many themselves. Thus, in order to understand the effectiveness of online support groups, a critical first step is to understand the factors that influence members' decisions to remain in them.

Maintaining membership in an online (or offline) group is a fundamental component of commitment to that group [2, 4 15]. In their model of group socialization, Levine and Moreland analyzed the antecedents and consequences of

Introduction

ONLINE SUPPORT GROUPS

Large Number of Online Health Support Groups

28% OF INTERNET USERS (Horrigan, 2001)

76% of people joined online health groups for emotional and informational support

(Ridings & Gefen, 2004)

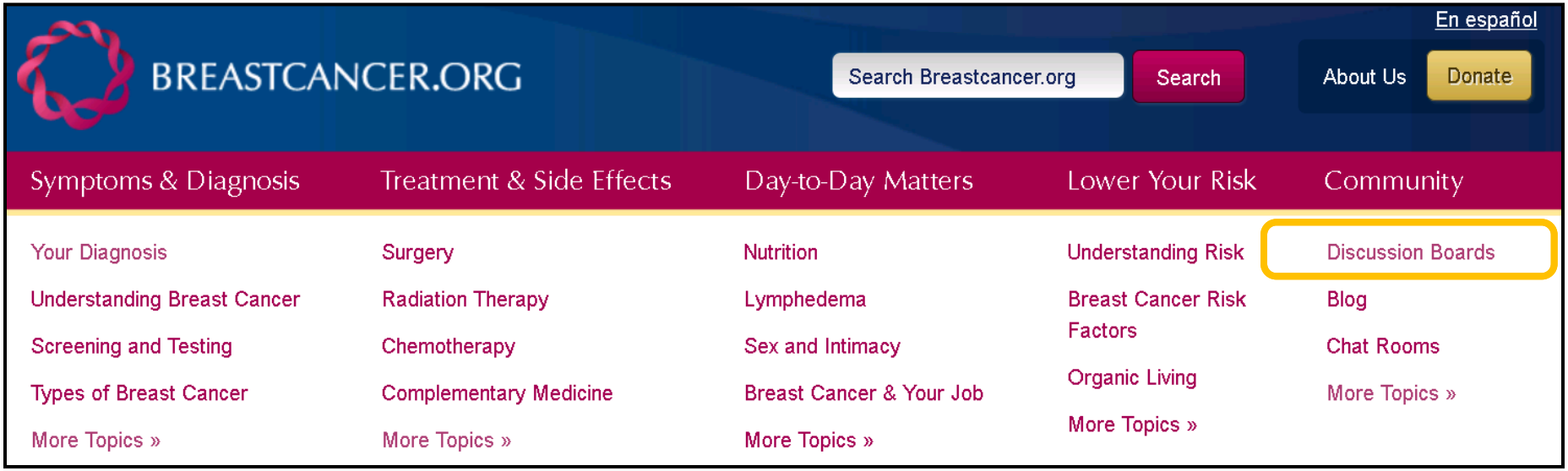
Two Types of Social Support

- Emotional support provides understanding, encouragement, sympathy, affirmation or caring

- Informational support provides advice, referrals or knowledge

(Bambina, 2007)

Research Site



Oct 2001 to Nov 2011

90,242 users

1,562,459 posts

68,158 discussion threads

Examples of Emotional Support

Two years ago yesterday my doctor called me and said "We made a mistake-you have breast cancer“ …

……

But here I am 2 years later..I am in the best shape of my life. My company has grown in ways I never imagined. I have received blessing after blessing. I have traveld to Hawaii, Costa Rica, San Francisco, NYC, SC, all since chemo ened.

Examples of Emotional Support

Two years ago yesterday my doctor called me and said "We made a mistake- you have breast cancer“ …
……

Yep, you are doing amazing!
See, all of us that told you "it does get better" at the beginning were right!

Congratulations on 2 years, what a great accomplishment! I hope you have many, many more years cancer-free.

::::::clapping.....dancing....tears in my eyes.....::::::so proud of you and your family.
awesome tiffany. there are no words to describe the power, truth and strength pouring out of you
blessings and love***

Examples of Informational Support

……
Help I am no expert..I have been tossed into this world of cancer and I am supposed to decide on chemo or not?? Talk about being in a boat w/out a paddle.....HELP!

Examples of Informational Support

……
Help I am no expert..I have been tossed into this world of cancer and I am supposed to decide on chemo or not?? Talk about being in a boat w/out a paddle.....HELP!

Has your doctor tested your tumor for the Oncotype score? It can help give more information on the % benefit for chemo…

I had a very large DCIS tumor too with about 4mm of invasive cancer. I was told by two oncologists no chemo, 1 oncologist said yes to the chemo…

I would encourage you to get a second opinion, preferably from an NCI-designated comprehensive cancer center.

Study 1:

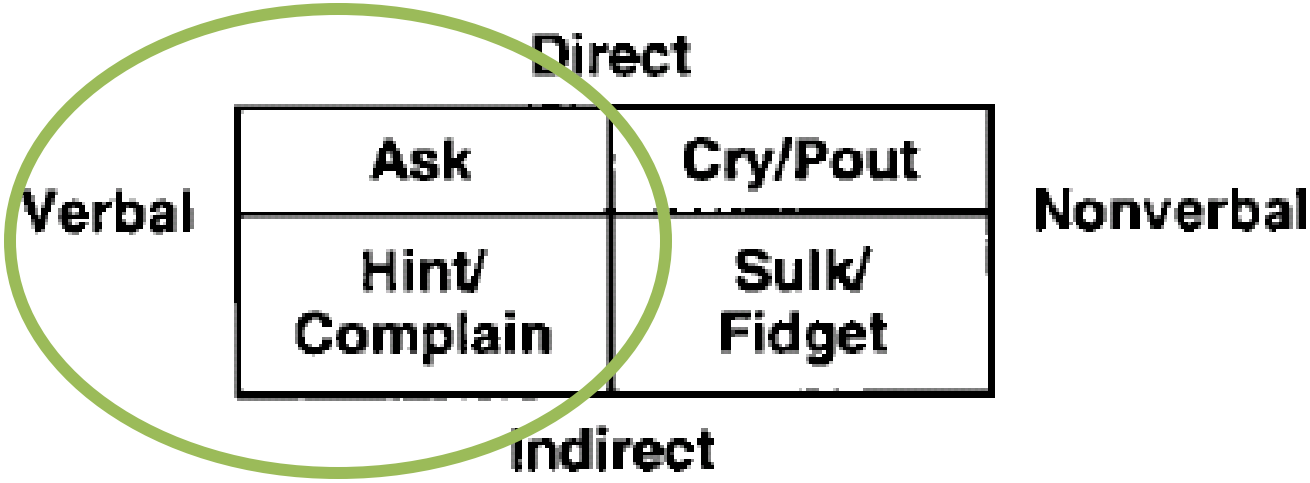
LANGUAGE STRATEGIES FOR SOCIAL SUPPORT RESPONSE

Social Support Communication

- A dynamic process (Pearlin & McCall, 1990)
  - Support elicitation
  - Support provision
  - Evaluation
- Little research examines the dynamic process as a whole
  - How to make effective social support exchange?

Strategies to Elicit Support

- Social support activation model (Barbee et al.,1993)
  - Support seeking behaviors can be categorized along two dimensions



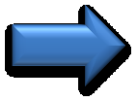
- In computer-mediated environments, the common applicable elicitation strategies are direct and indirect verbal.

Online Support Requests (Bambina, 2007)

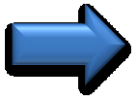
- Clear-cut requests: direct questions
- Negative sentiments: e.g., distress, fear, and sorrow
- Accounts: past experiences, prognosis, and updates

Research Question

In an online health support group

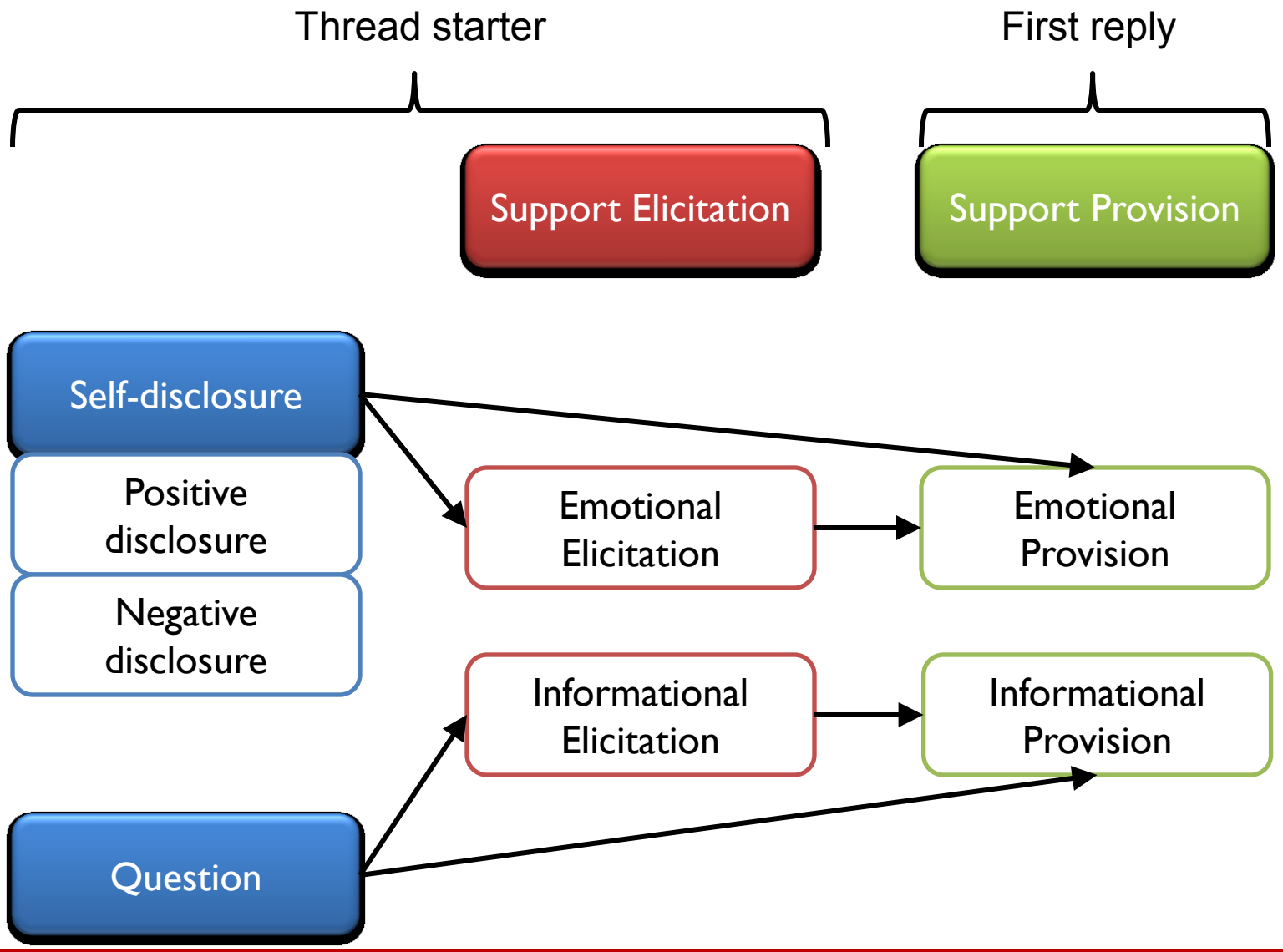


Emotional support



Informational support

Hypothesized Model



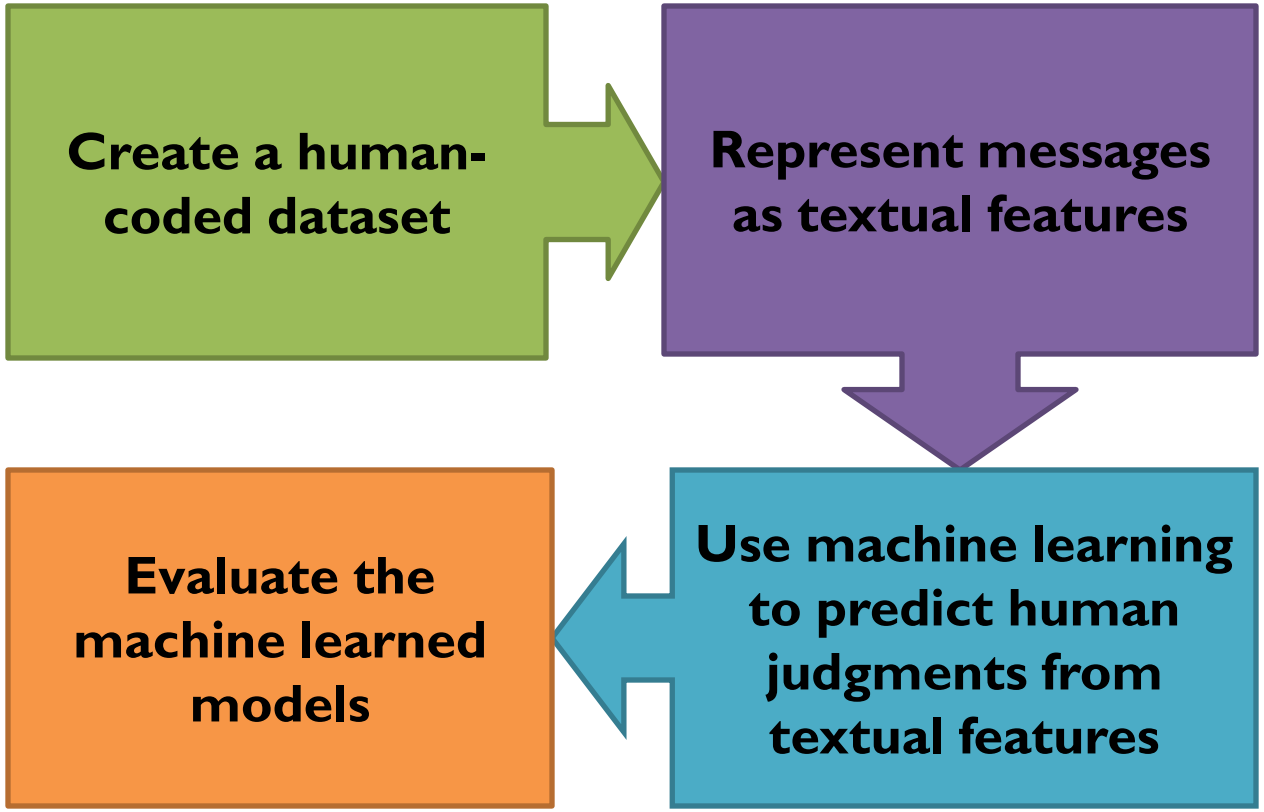
Phases of Study

Develop automated measures of self-disclosure, question, support elicitation, and provision

Examine the model using both hand-coded data and machine-coded data

Process of Develop Automatic Measurements

Machine Learning Approach



Create the Human-Coded Dataset

amazonmechanical turk
Artificial Artificial Intelligence

Your Account | HITs | Qualifications

Already have an account?
Sign in as a Worker | Requester

Introduction | Dashboard | Status | Account Settings

Mechanical Turk is a marketplace for work.
We give businesses and developers access to an on-demand, scalable workforce.
Workers select from thousands of tasks and work whenever it's convenient.

200,397 HITs available. View them now.

Make Money
by working on HITs

HITs - *Human Intelligence Tasks* - are individual tasks that you work on. Find HITs now.

As a Mechanical Turk Worker you:

- Can work from home
- Choose your own work hours
- Get paid for doing good work

Find an interesting task → Work → Earn money

Find HITs Now

or learn more about being a **Worker**

Get Results
from Mechanical Turk Workers

Ask workers to complete HITs - *Human Intelligence Tasks* - and get results using Mechanical Turk. Register Now

As a Mechanical Turk Requester you:

- Have access to a global, on-demand, 24 x 7 workforce
- Get thousands of HITs completed in minutes
- Pay only when you're satisfied with the results

Fund your account → Load your tasks → Get results

Get Started

NOTE!

- **This task is for native English speakers.**
- **Only serious responses will be paid. Thank you for your understanding.**

Instructions

- You will be reading messages posted to several online cancer support groups. Please indicate the extent to which each message is providing social support (emotional support and/or information support) to other participants in the group.
- For the people who use IE as your web browser, if you can't see the messages in the second section, please reload the page and hit the "NO" button when the security warning dialogue box shows up.

(1) Read the definition of social support.

- There are two kinds of social support, **emotional support and informational support**.
- **Emotional support** messages provide understanding, encouragement, affirmation, sympathy, or caring.
- **Informational support** messages provide advice, referrals or knowledge.
- To see the definitions and examples of emotional and informational support, please follow this link.

(2) Please do the following:

- Read the subject line and message body highlighted in yellow.
- Answer the questions associated with each message.
- We've highlighted the subject line and message body in yellow. The quoted material in the message is included as context. Please make your judgement based on the highlighted material.

Please carefully read the highlighted message:

Subject: Mixed Reviews

My mom was diagnosed with Ductal Carcinoma In Situ with extensive Comedonecrosis. She has had two lumpectomies and the second lumpectomy report is saying that it may be invasive. The surgeon thought that he had clear margins. They want to take out a lymph node just to make sure that it hasn't spread, so my question is, is it non-invasive or invasive? Aaarrrrggghhh, it is so aggravating and stressful to get contradicting reports.

Can you get a copy of the path report from the 2nd lumpectomy? It should clearly state the impression, whether DCIS only or both DCIS and IDC.

How much **emotional support** does this message provide?

- 1 (none)
- 2

Create the Human-Coded Dataset

amazonmechanical turk
Artificial Artificial Intelligence

Thread Starter

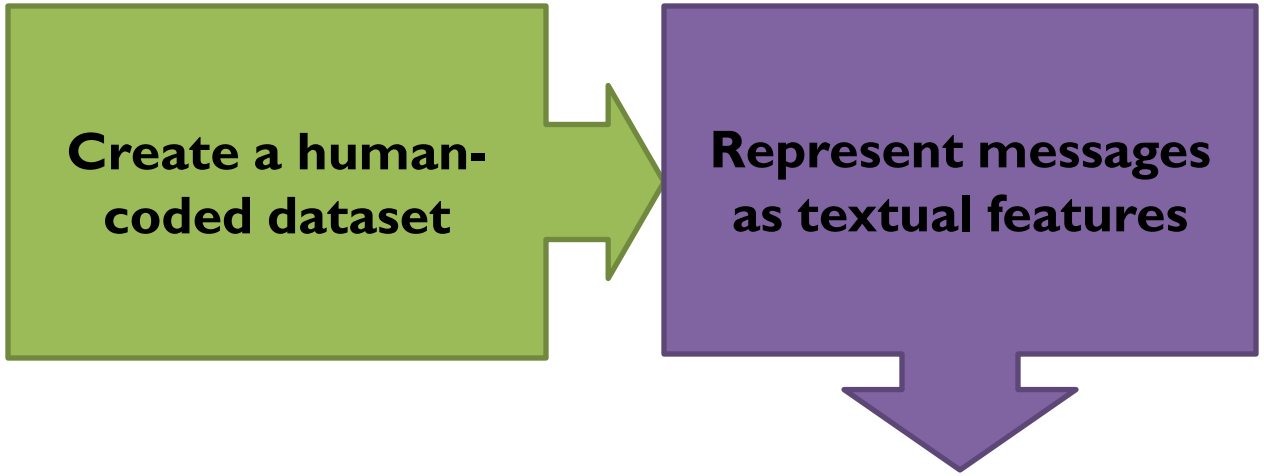
- Self-disclosure: To what extent does this message contain positive/negative self-disclosure?
  - ICC(positive self-disclosure) = 0.90
  - ICC(negative self-disclosure) = 0.94
- Question: To what extent is this message asking a question?
  - ICC(question) = 0.91
- Elicitation: To what extent is this message seeking emotional/informational support?
  - ICC(emotional support elicitation) = 0.91
  - ICC(informational support elicitation) = 0.95

First reply

- Provision: How much emotional/informational support does this message provide?
  - ICC(emotional support provision) = 0.92
  - ICC(informational support provision) = 0.92

Automatically Measure Social Support

Machine Learning Approach



Three Types of Textual Features

Type	Sample Feature	Vocabulary/Pattern
Generic dictionaries (LIWC)	We	We, our, us
Topic modeling	Post-surgery problems	Pain, blood, tamoxifen, symptom
	Spiritual	Love, god, prayer, bless, peace
Linguistics Pattern	Advice	<if+you>

If you are having pain that is bad enough that you are fearing a deadly disease, then maybe your doctor needs to actually diagnose where this pain is coming from.
God bless, and we will help each other all we can.

Three Types of Textual Features

Type	Sample Feature	Vocabulary/Pattern
Generic dictionaries (LIWC)	We	We, our, us
Topic modeling	Post-surgery problems	Pain, blood, tamoxifen, symptom
	Spiritual	Love, god, prayer, bless, peace
Linguistics Pattern	Advice	<if+you>

If you are having pain that is bad enough that you are fearing a deadly disease, then maybe your doctor needs to actually diagnose where this pain is coming from.
God bless, and we will help each other all we can.

Three Types of Textual Features

Type	Sample Feature	Vocabulary/Pattern
Generic dictionaries (LIWC)	We	We, our, us
Topic modeling	Post-surgery problems	Pain, blood, tamoxifen, symptom
	Spiritual	Love, god, prayer, bless, peace
Linguistics Pattern	Advice	<if+you>

If you are having pain that is bad enough that you are fearing a deadly disease, then maybe your doctor needs to actually diagnose where this pain is coming from.
God bless, and we will help each other all we can.

Three Types of Textual Features

Type	Sample Feature	Vocabulary/Pattern
Generic dictionaries (LIWC)	We	We, our, us
Topic modeling	Post-surgery problems	Pain, blood, tamoxifen, symptom
	Spiritual	Love, god, prayer, bless, peace
Linguistics Pattern	Advice	<if+you>

If you are having pain that is bad enough that you are fearing a deadly disease, then maybe your doctor needs to actually diagnose where this pain is coming from. God bless, and we will help each other all we can.

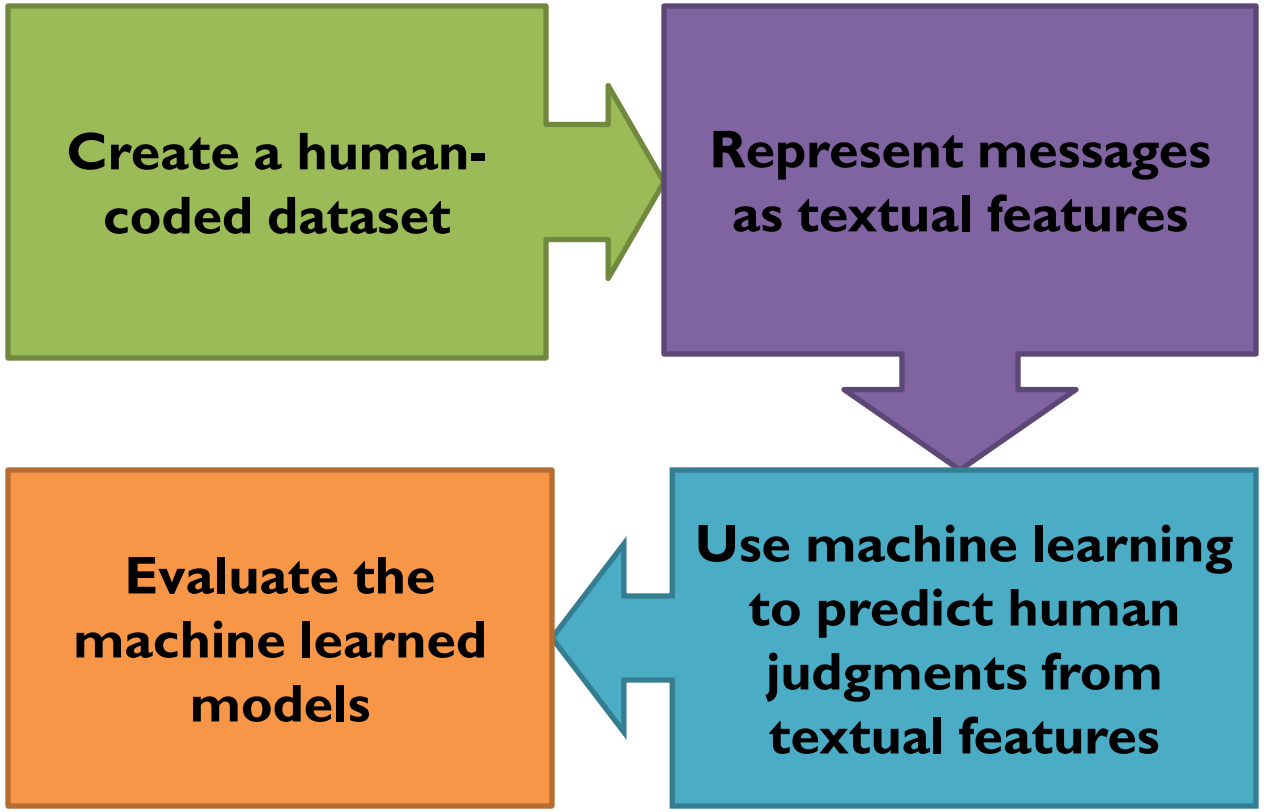
Three Types of Textual Features

Type	Sample Feature	Vocabulary/Pattern
Generic dictionaries (LIWC)	We	We, our, us
Topic modeling	Post-surgery problems	Pain, blood, tamoxifen, symptom
	Spiritual	Love, god, prayer, bless, peace
Linguistics Pattern	Advice	<if+you>

If you are having pain that is bad enough that you are fearing a deadly disease, then maybe your doctor needs to actually diagnose where this pain is coming from.
God bless, and we will help each other all we can.

Automatically Measure Social Support

Machine Learning Approach



Construct Machine Learning Models

Social Support Concept	Correlation
Positive self-disclosure	0.44
Negative self-disclosure	0.59
Question	0.78
Emotional support elicitation	0.59
Informational support elicitation	0.75
Emotional support provision	0.81
Informational support provision	0.85

Important Features in Machine Learning Models

Support	Emotional			Informational		
$R_{ModelXHuman}$	0.76			0.80		
Most Important Features	+	0.5948	Sentence count	+	0.8918	Sentence count
	+	0.4531	Emotional support	+	0.4084	Word count per sentence
	+	0.4196	We	-	0.2391	Sentiment_SubjStrong_PerWord
	+	0.3971	You	+	0.2331	<if+you>
	-	0.3703	She/he	-	0.2163	I
	+	0.2833	Spiritual	-	0.2114	Spiritual
	-	0.2642	<if+you>	-	0.1999	We
	-	0.2611	Positive life events	+	0.1998	Post-surgery problems
	+	0.2439	Adjusting to diagnosis	-	0.167	Forum communication
	-	0.2398	Time	-	0.1639	Religion

Important Features in Machine Learning Models

Support	Emotional			Informational		
$R_{ModelXHuman}$	0.76			0.80		
Most Important Features	+	0.5948	Sentence count	+	0.8918	Sentence count
	+	0.4531	Emotional support	+	0.4084	Word count per sentence
	+	0.4196	We	-	0.2391	Sentiment_SubjStrong_PerWord
	+	0.3971	You	+	0.2331	<if+you>
	-	0.3703	She/he	-	0.2163	I
	+	0.2833	Spiritual	-	0.2114	Spiritual
	-	0.2642	<if+you>	-	0.1999	We
	-	0.2611	Positive life events	+	0.1998	Post-surgery problems
	+	0.2439	Adjusting to diagnosis	-	0.167	Forum communication
	-	0.2398	Time	-	0.1639	Religion

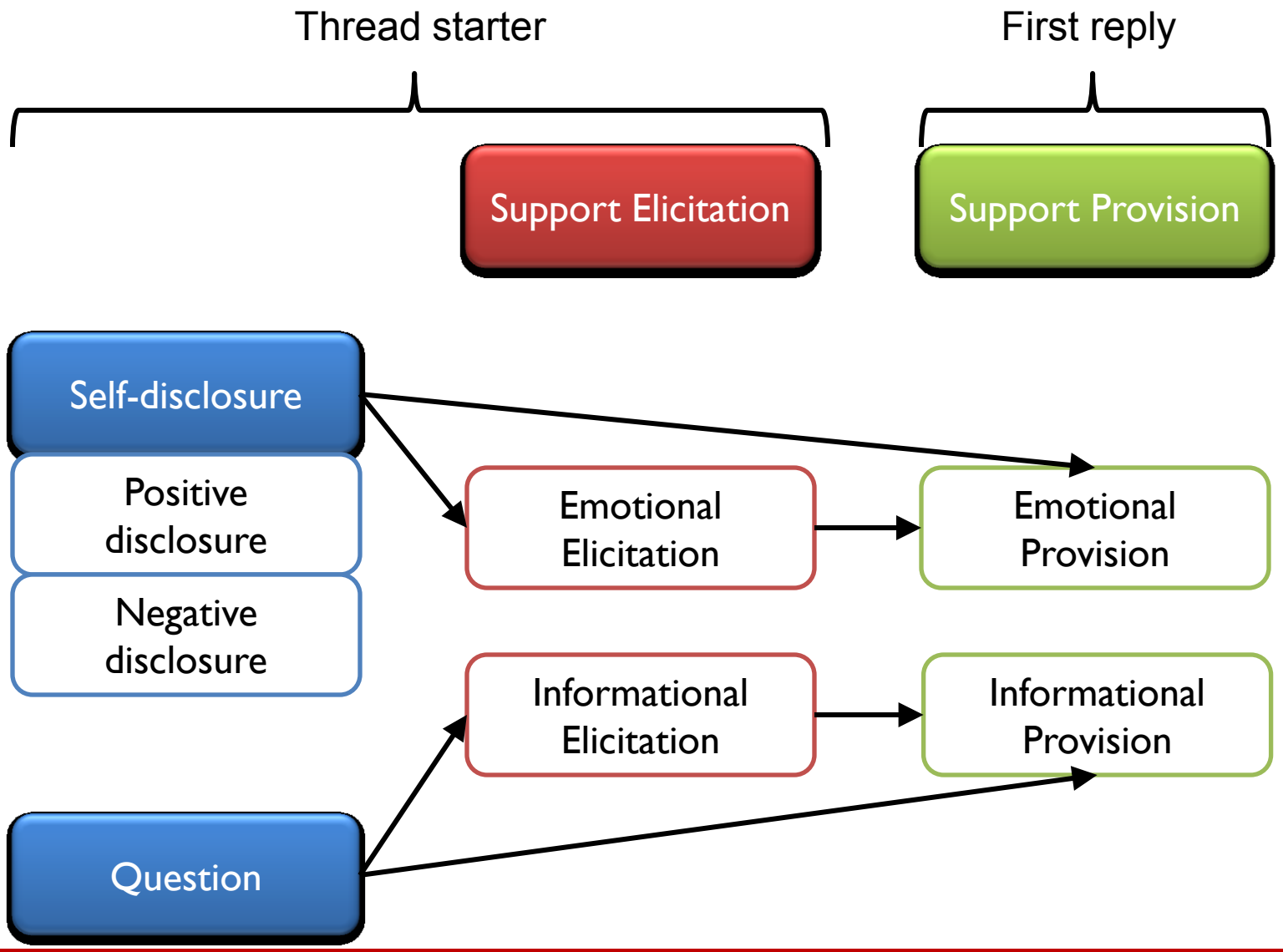
Important Features in Machine Learning Models

Support	Emotional			Informational		
$R_{ModelXHuman}$	0.76			0.80		
Most Important Features	+	0.5948	Sentence count	+	0.8918	Sentence count
	+	0.4531	Emotional support	+	0.4084	Word count per sentence
	+	0.4196	We	-	0.2391	Sentiment_SubjStrong_PerWord
	+	0.3971	You	+	0.2331	<if+you>
	-	0.3703	She/he	-	0.2163	I
	+	0.2833	Spiritual	-	0.2114	Spiritual
	-	0.2642	<if+you>	-	0.1999	We
	-	0.2611	Positive life events	+	0.1998	Post-surgery problems
	+	0.2439	Adjusting to diagnosis	-	0.167	Forum communication
	-	0.2398	Time	-	0.1639	Religion

Important Features in Machine Learning Models

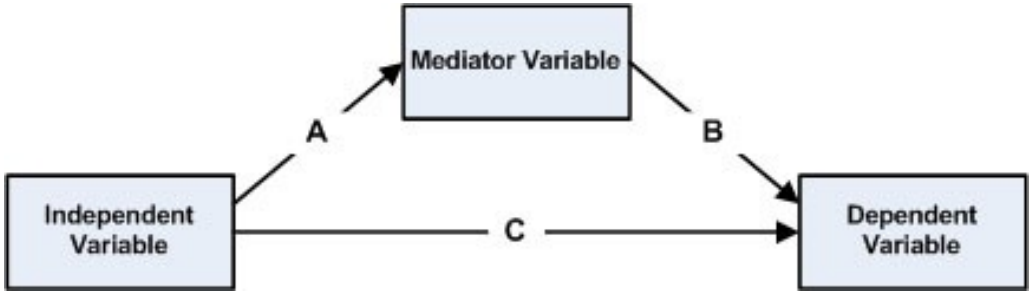
Support	Emotional			Informational		
$R_{ModelXHuman}$	0.76			0.80		
Most Important Features	+	0.5948	Sentence count	+	0.8918	Sentence count
	+	0.4531	Emotional support	+	0.4084	Word count per sentence
	+	0.4196	We	-	0.2391	Sentiment_SubjStrong_PerWord
	+	0.3971	You	+	0.2331	<if+you>
	-	0.3703	She/he	-	0.2163	I
	+	0.2833	Spiritual	-	0.2114	Spiritual
	-	0.2642	<if+you>	-	0.1999	We
	-	0.2611	Positive life events	+	0.1998	Post-surgery problems
	+	0.2439	Adjusting to diagnosis	-	0.167	Forum communication
	-	0.2398	Time	-	0.1639	Religion

Hypothesized Model



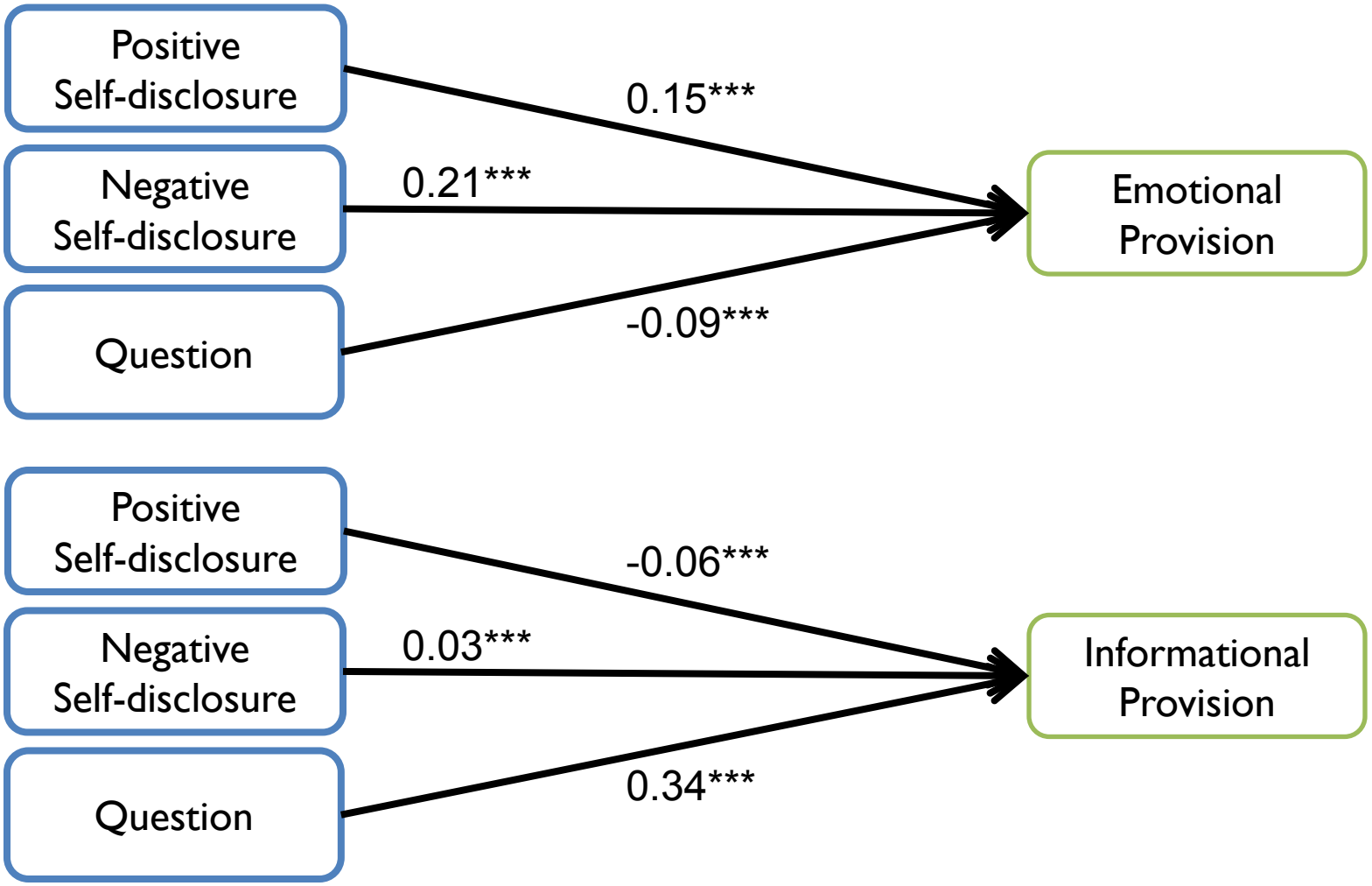
Regression and Mediation Analysis

- Regression analysis: $Y = \alpha + \beta_1 X_1 + \ldots + \beta_k X_k + \varepsilon$
  - X: Independent variable (feature)
  - Y: Dependent variable (label)
  - $\alpha, \beta_1, \ldots, \beta_k$: Coefficients
    - * marks significance level
- Mediation analysis

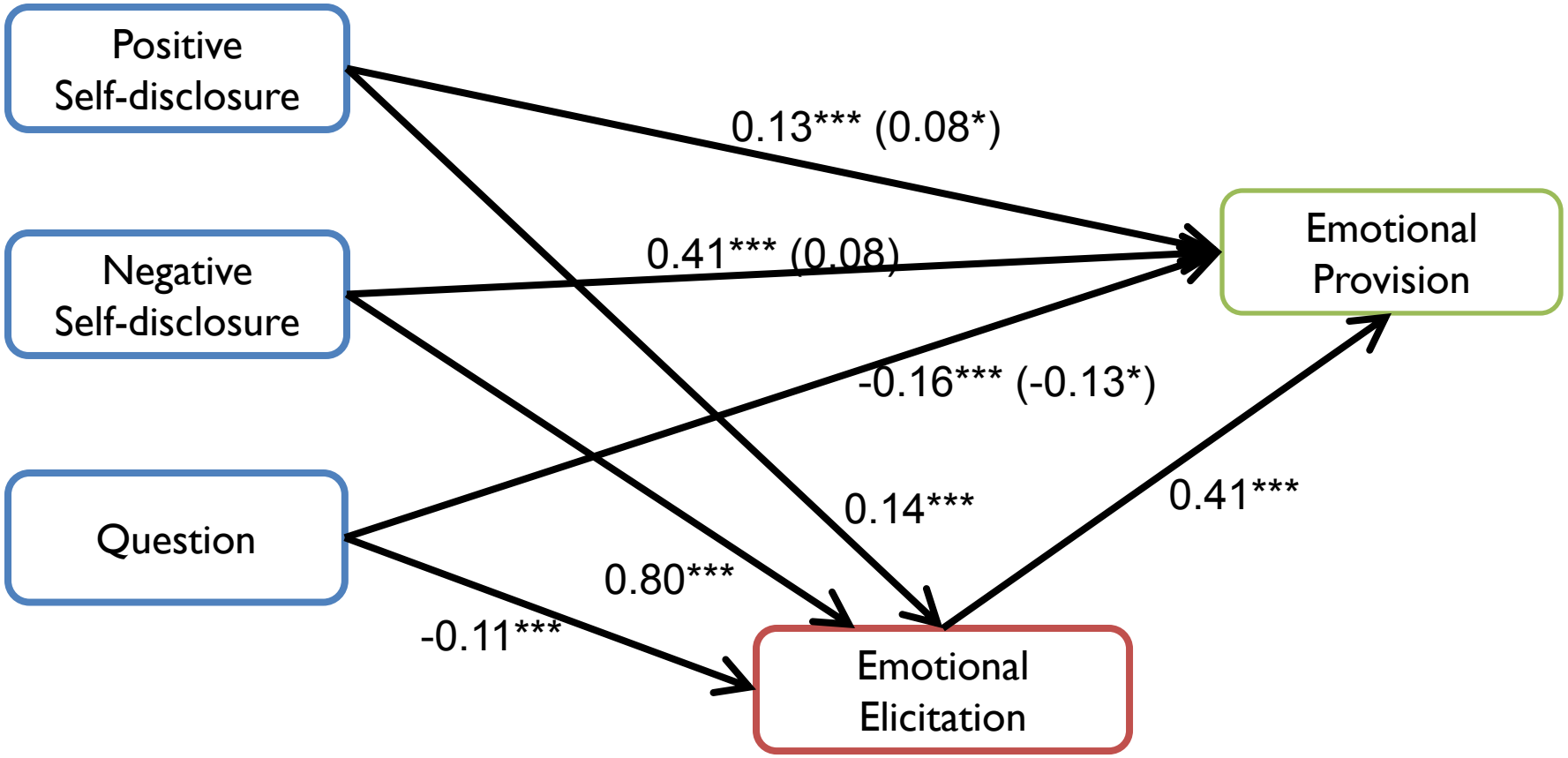


(from Wikipedia)

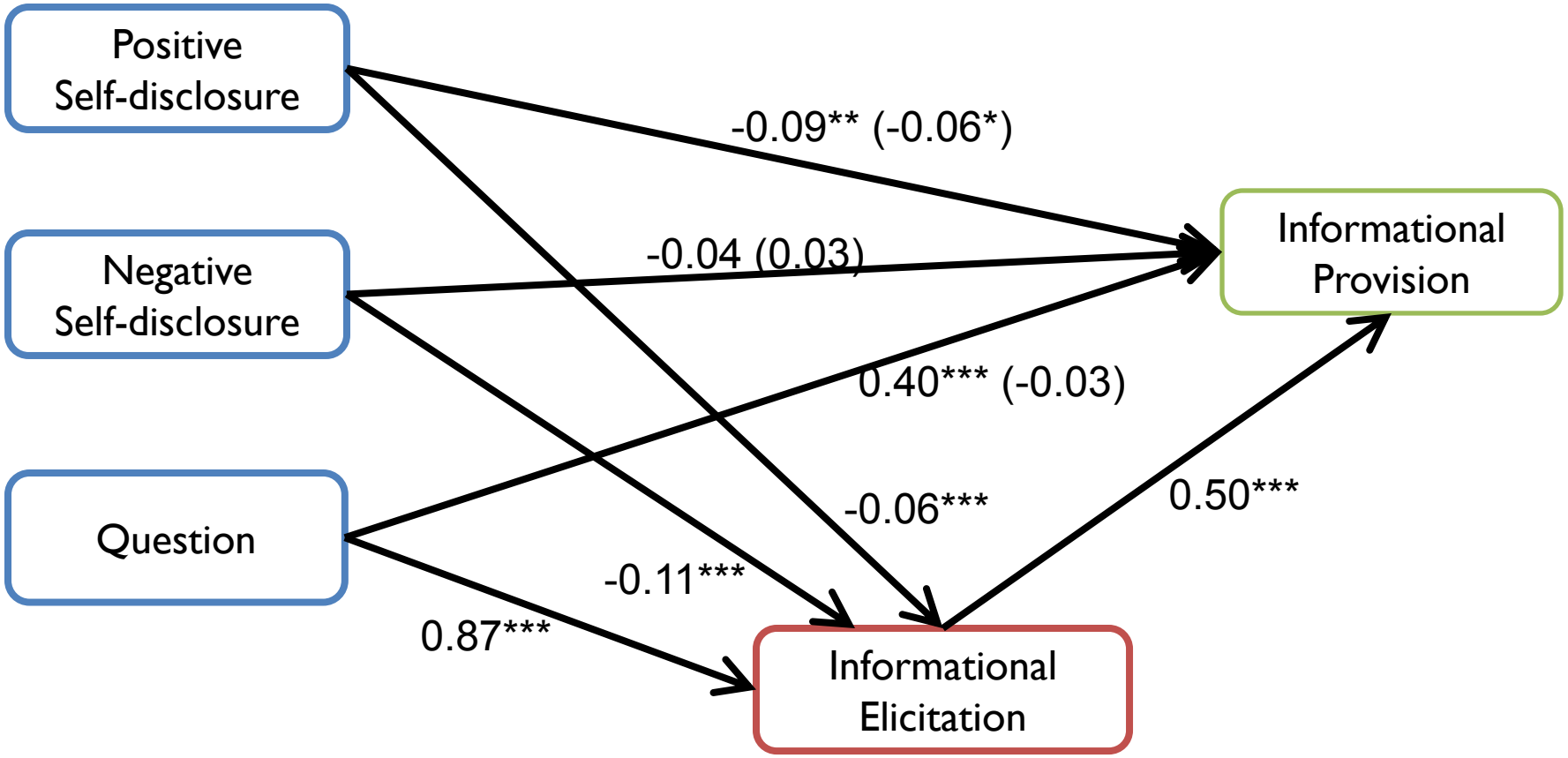
Regression Analysis Using Machine-Coded Data



Mediation Analysis Using Human-Coded Data



Mediation Analysis Using Human-Coded Data



Summary

- The significantly positive effects of self-disclosure and question in machine models are consistent with human models.

- Analysis based on human-coded data has confirmed our hypothesized model

Study 2:

SOCIAL SUPPORT & MEMBERSHIP COMMITMENT

Commitment is Important

Commitment is defined as continued participation

- Individual participant
  - People can only receive benefits by participating in the online conversations
- Group
  - Participation is essential to provide benefits to others

Which kind of support would keep people around longer?

Emotional Support

You have had such a difficult road, but you still manage to do well. I am truly inspired by you. Big hugs :>

Informational Support

Extranodal extension occurs when the tumor extends through the wall of the lymph node.

Hypotheses

In an online health support group

Receiving emotional support

Duration of participation

Receiving informational support

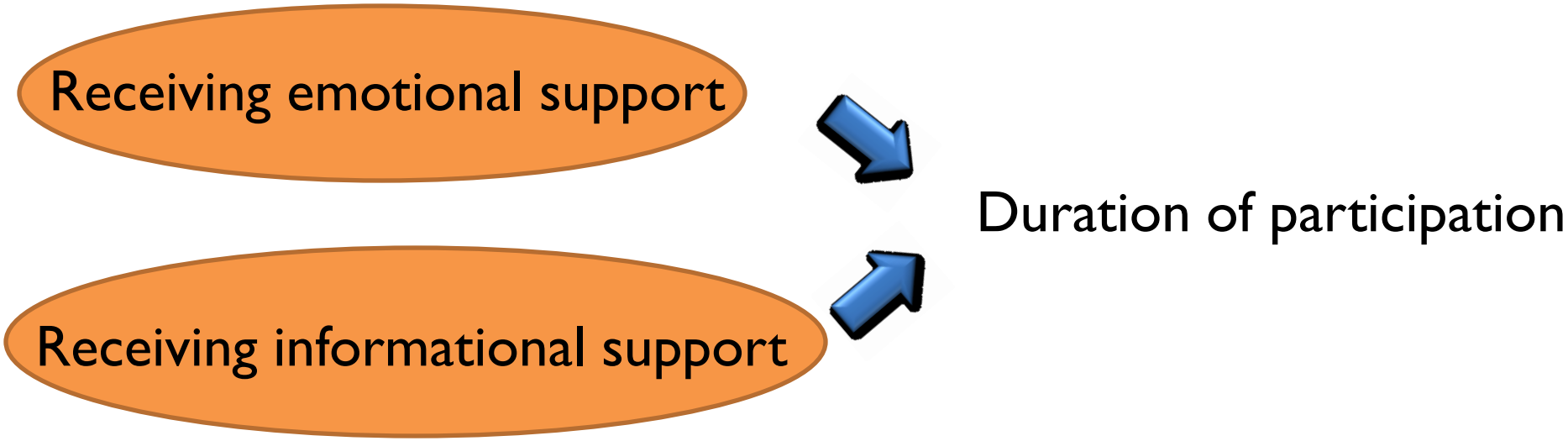
Phases of Study

Develop automated measures of emotional and informational support

Examine the relationship between social support and duration of participation

Social Support & Commitment

In an online health support group



Assuming that participants are exposed to all messages posted to threads during the week they also posted

Survival Analysis

- A statistical technique for investigating time-related outcomes, such as whether or when an event occurs
  - Time until a member leaves the group

- 10-year longitudinal data

Variables for Survival Analysis

IVs: Receiving social support in a week

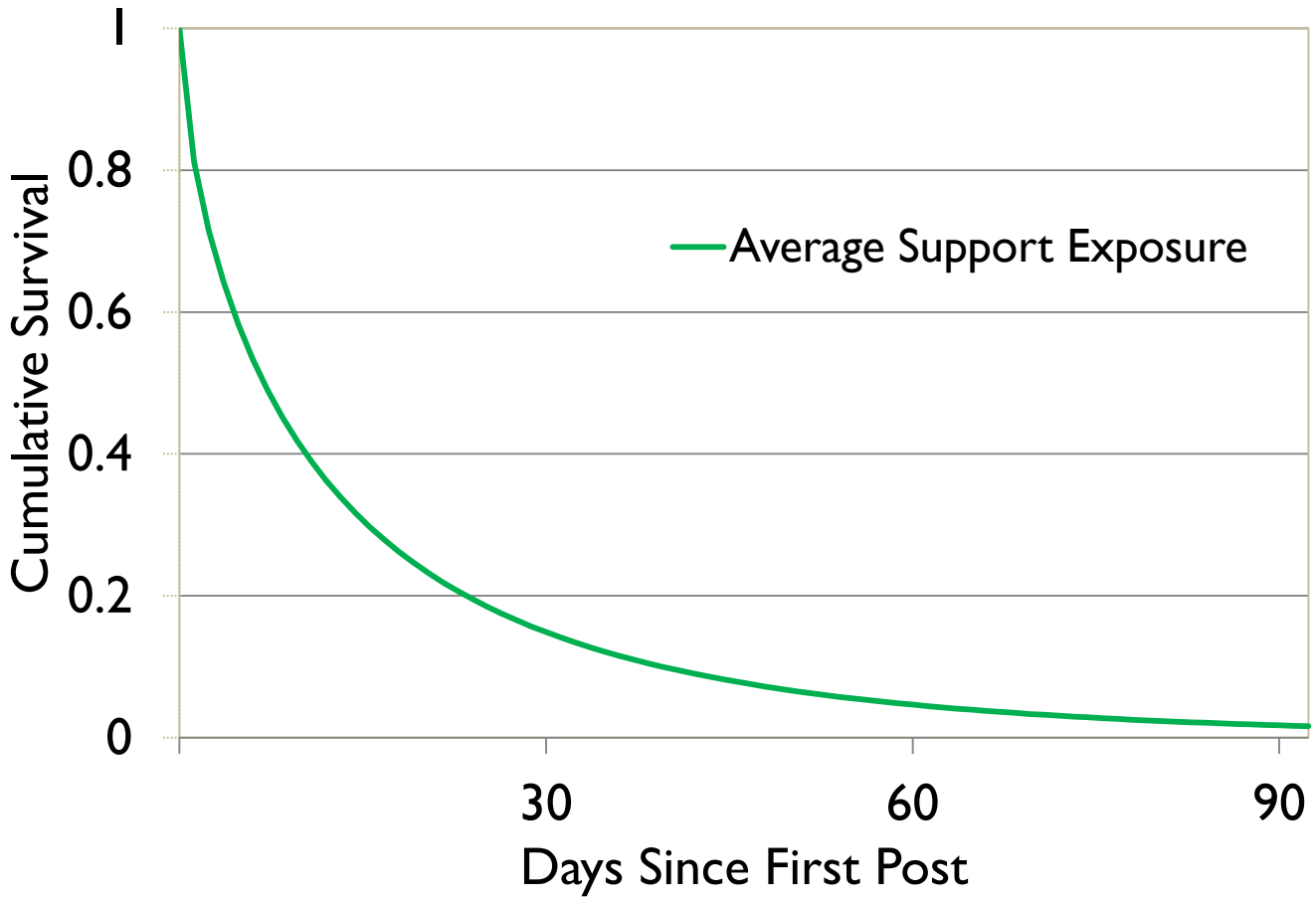
- Post Count Exposure
- Emotional Support Exposure
- Informational Support Exposure
- Post Count Exposure * Emotional Support Exposure
- Post Count Exposure * Informational Support Exposure

DV: Dropout

- If a user fails to post again in the next 12 weeks

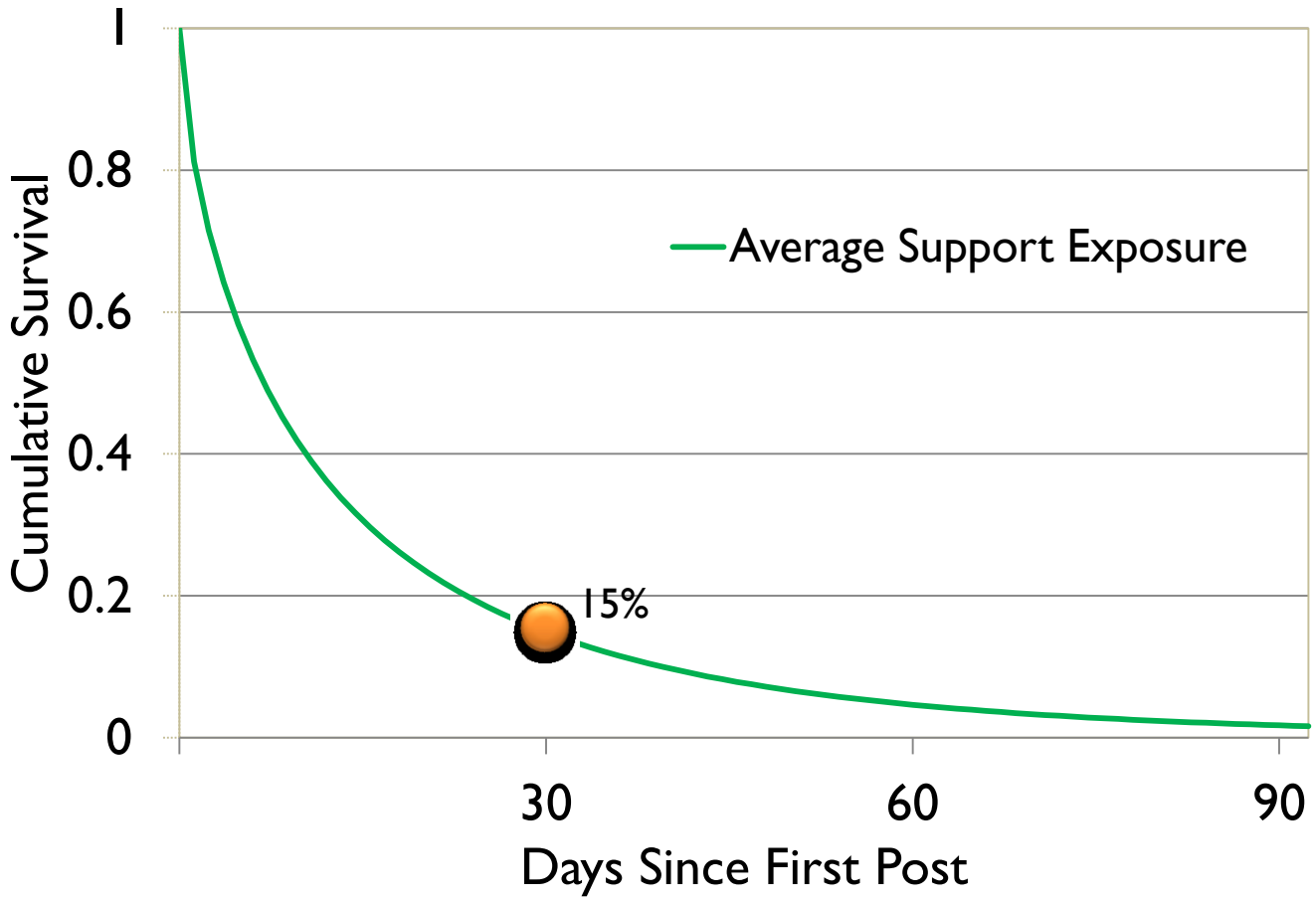
Survival Curves

- Plot percent survival as a function of time



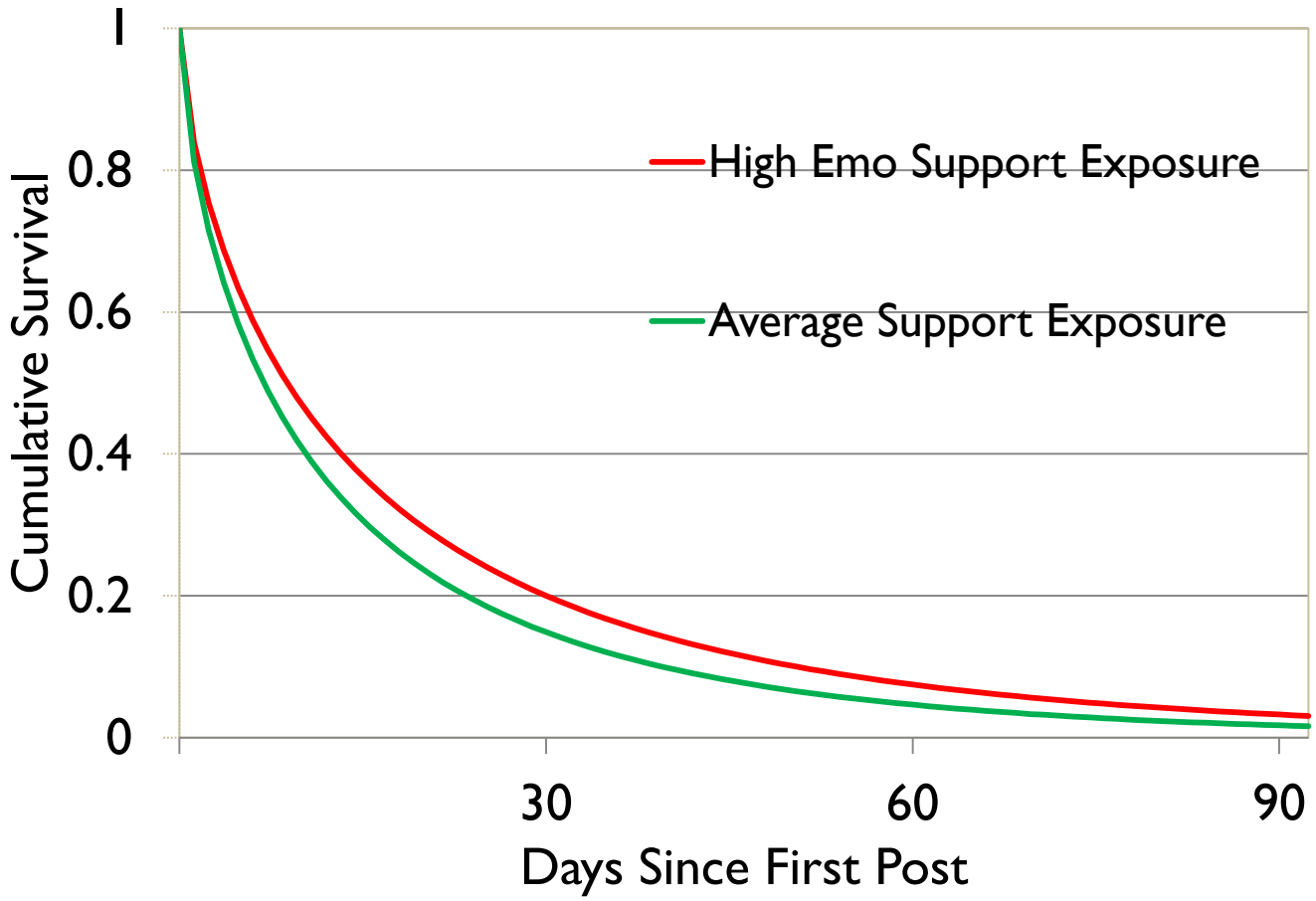
Survival Curves

- People dropout relatively quickly in this group



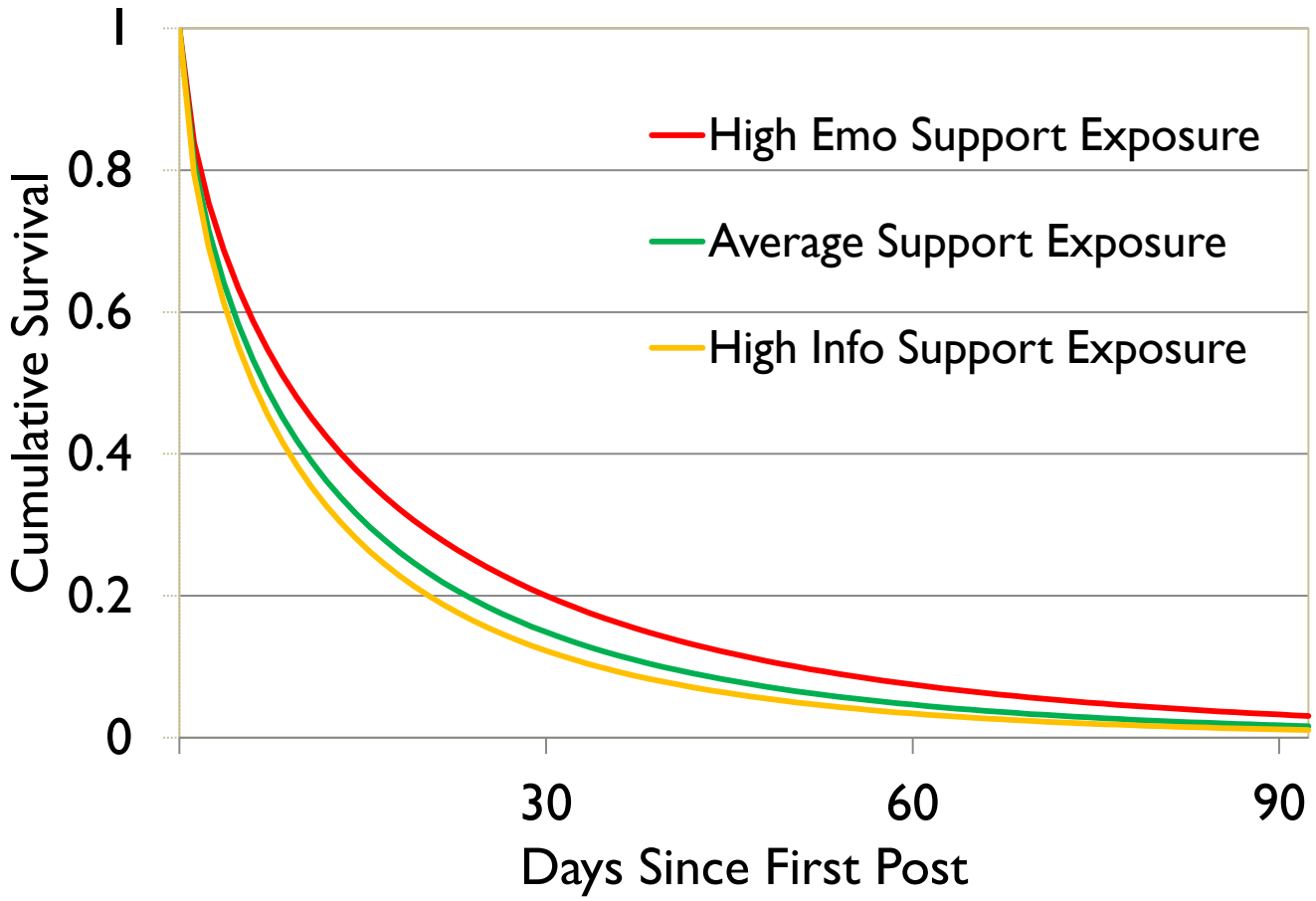
High Emotional Support Exposure

- 16% more likely to remain in the group



High Informational Support Exposure

- 10% more likely to leave the group



Discussion

- Emotional and informational support have different effects
  - Quality of resource
    - "Good" emotional support vs "poor" informational support
  - Time needed to obtain resource
    - Long-term social relationship is necessary for emotional support
    - Quick interactions are sufficient for informational support
  - Pre-existing differences among recipients
    - Emotional support seekers also look for relationships
    - Informational support seekers don't care about relationships

Limitations

- Although we use longitudinal data, our findings are correlational
  - Random-assignment experiments
- No direct measures of reading behavior

Contributions and Implications

- Methodological: We obtained automated coding of social support that have comparable quality to human judgments
  - Easy to generalize
  - Automatic content analysis and intervention in online support groups

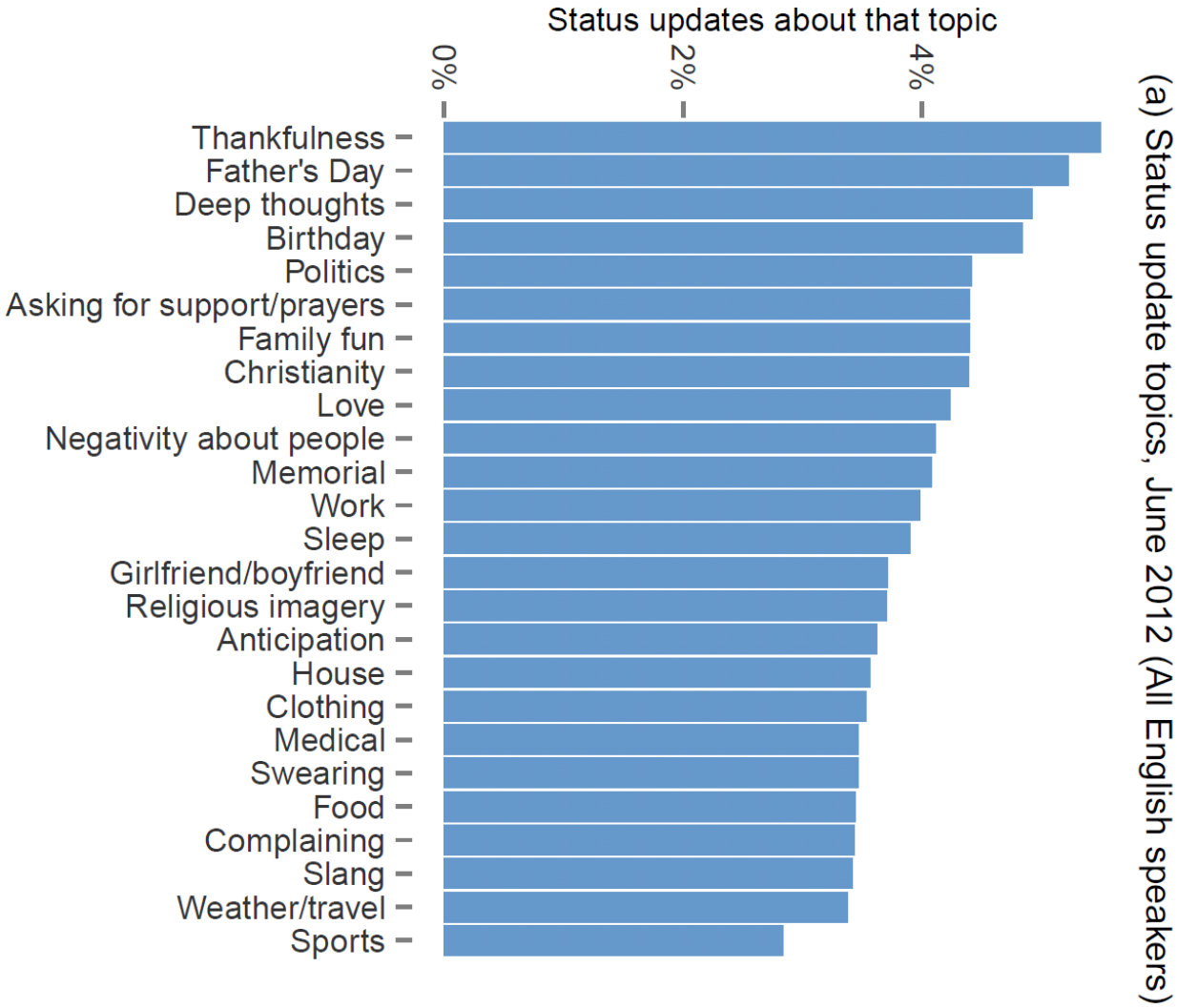
- Substantive
  - Different strategies to acquire emotional & informational support
  - Emotional support & informational support were differentially associated with commitment

Generalization

- Different uses of language are associated with different response rate

- Responses keep people around and build their commitment to the group
  - Online dating sites
  - Social networking sites
  - …

What are people talking about in status updates?

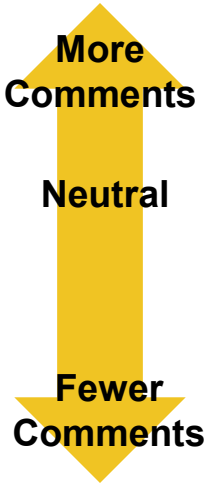


facebook

- t = 1
  - >=3 ngrams in t dictionary
- 50% of status updates have 2 topics

		#. Comments [1]	
		Estimate	Est. Comments
	(Intercept)	0.311***	1.04
Controls	Male	-0.006***	1.02
	Age (years) [3]	0.030***	1.19
	Days since registration [3]	0.010***	1.09
	Friend count [2]	-0.048***	0.83
	Comments per prev. post [2]	0.066***	1.38
	IsWeekday	0.045***	1.27
	Post views [2]	0.134***	1.78
	Post length (words) [2]	0.025***	1.17
Topics	Medical	0.084***	1.48
	Swearing	0.028***	1.18
	House	0.028***	1.18
	Asking for support/prayers	0.026***	1.17
	Deep thoughts	-0.031***	0.90
	Family fun	-0.032***	0.90
	Religious imagery	-0.042***	0.86
	Love	-0.047***	0.83
	Sleep	-0.047***	0.83
	Christianity	-0.069***	0.75

1: Logged (base 10) 2: Logged (base 10), standardized, centered
3: standardized and centered. *p<0.05, **p<0.01, ***p<0.001



Question?

Yi-Chia Wang (yichiaw@cs.cmu.edu)