

# How social networks help college students trying to lose weight: Analyzing the Online Conversation

Lilla Orr

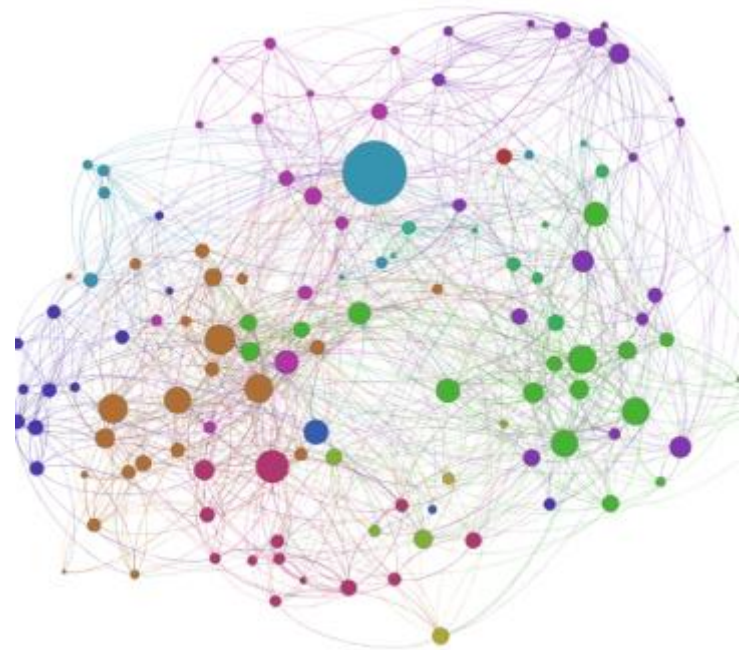
BS Biochemistry, Doctoral Student Political Science

Authors: Gina Merchant, James H. Fowler, Guadalupe X. Ayala, Linda C. Gallo, Jim D. Hollan, Kevin Patrick

# Social networks

A web of social relationships and their corresponding properties (Glanz, 2008; Leroux, 2013)

- Exchange of social support and social capital
- Normative influence



# Social network influence

## Social capital

- **Resources: called upon or simply available if needed**
- Difficult to measure; limited agreement on definition

## Normative influence

- **We are not necessarily aware of it operating**
- Hard to measure; not usually measured correctly

## Social support

- **A transaction that is real or perceived**
- Most commonly measured

[illegible]

**“I mean, we all do it for the likes.”**  
- *Project SMART study participant*

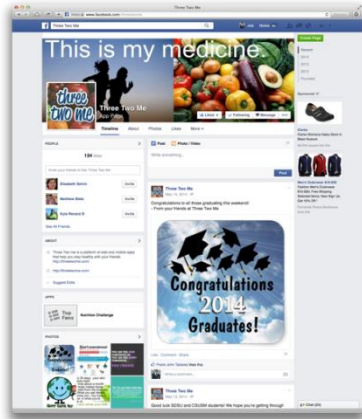
- 1.5 billion use Facebook each month
- 4 billion pieces of content shared each day
- Dynamic activity (sharing, feedback, exposure)



somee cards  
user card

# Data source: SMART Intervention

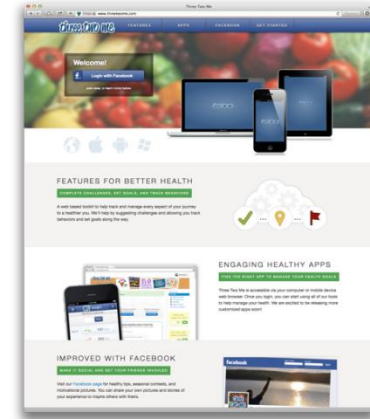
## 1) Facebook



## 2) Apps



## 3) Website



## 4) Texts



## 5) Email



## 6) Health coach



# Participants

**Table 1. Participant characteristics by intervention group (N = 329)**

	Total (N = 329)	Control (n = 167)	Treatment (n = 162)
Age (years), mean (SD)	22.6 (3.8)	22.7 (3.8)	22.4 (3.7)
Race, n (%)			
White	135 (41.0)	66 (39.5)	69 (42.6)
Other/Multiple	96 (29.2)	51 (310.54)	45 (27.8)
Asian	80 (24.3)	41 (24.6)	39 (24.1)
Black	12 (3.6)	7 (4.2)	5 (3.1)
American Indian/Alaskan/Pacific Islander	6 (1.8)	2 (1.2)	4 (2.5)
Ethnicity, n (%)			
Hispanic	103 (31.3)	54 (67.7)	49 (30.4)
Undergraduate (yes), n (%)	162 (49.2)	87 (52.1)	75 (46.3)
Anthropometrics, mean (SD)			
Body mass index (BMI)	28.9 (2.8)	28.9 (2.7)	28.9 (2.9)
Waist circumference (cm)	87.4 (8.9)	87.6 (8.8)	87.3 (8.9)



# ThreeTwoMe



# Does being in a weight-loss trial affect how much you talk about healthy living with your online social network?

$H_1$  Compared to control participants, treatment participants post more health-related content in their Facebook status updates after joining the study.

$H_2$  Compared to those less engaged, treatment participants who are more engaged with the *ThreeTwoMe* page will post more health-related content in their Facebook status updates.



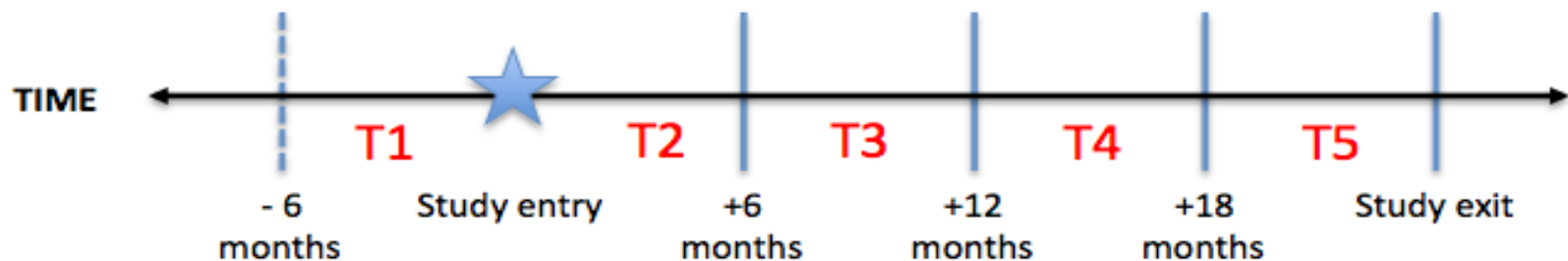
# Plan of analysis

## Supervised approach

Create a Healthy Active Lifestyle (HAL) Dictionary using posts made by the health coach on the study's Facebook page

## Facebook data

Broadcasted posts from Facebook's social graph,  $N = 358$



## Unit of analysis

Fraction of participants' status updates that contain at least one HAL unigram over a 30 month period

# Dictionary Creation

## Words scraped from

- ThreeTwoMe posts
- USDA National Nutrient Database
- Compendium of Physical Activity



Three Two Me

September 23, 2013

on WEDNESDAY we will start a 7 DAYS OF FITNESS as a "reset" button for all of us! we did this a little while back and it was a success – are you in?! Your Health Coach will be doing it right along with you and blogging along the way – Tell us you're in below!

Like · Comment · Share



4

118 people saw this post

Boost Post

## Inclusion Criteria

- Purposeful physical activity, healthy food
- Unigrams
- All grammatical forms of a root word
- Expert consensus

# Dictionary Evaluation

## Validity check on HAL dictionary

- Random sample ( $n = 2,614$ ) 5% of baseline posts
- Two researchers independently code status updates as HAL or non-HAL, reconcile differences

Does the post describe the poster engaging in  
past/current/planned purposeful...

physical activity/exercise?

dietary choices which we would consider part of  
a healthy active lifestyle?

# Validity results: Human coding

## Not HAL, but computer classified as HAL

“...let's just hug it out ok? APRIL FOOLS - IN 2 DAYS IM STOMPIN A MUDHOLE IN YOUR **FRUIT** LOOP PUNK ASS - TEAM BRING IT. -The Rock”

## Is HAL, but not computer classified as HAL

“Who wants to **hit legs** today?!”

## Questionable, computer classified as HAL

“Via Jen. My favorites are the "Land, HO!" and the "... Jesus".  
[http://tryphena.tumblr.com/post/5802996931/sylvysparrow-sofapizza-pleatedjeans-\*\*yoga\*\*](http://tryphena.tumblr.com/post/5802996931/sylvysparrow-sofapizza-pleatedjeans-yoga)”

# Validity results: Human coding

## Limitations

### + and – health behaviors

“After a week of **binge drinking** and eating out... I got 25 days to get ready 4 VEGAS!

Day 1: **cardio, chest, tri's, abs...**”

# Reliability and validity results

## Human coding reliability:

- Overall Kappa = 65%
- Diet Kappa = 75%
- Exercise Kappa = 62%

## Diagnostic validity:

- **Sensitivity:** 55%
- **Specificity:** 98%

Dictionary  
classifier

Human truth		
	HAL	not HAL
HAL	36	41
not HAL	29	2508

**PROBLEM:** Dictionary misses a lot of true HAL posts



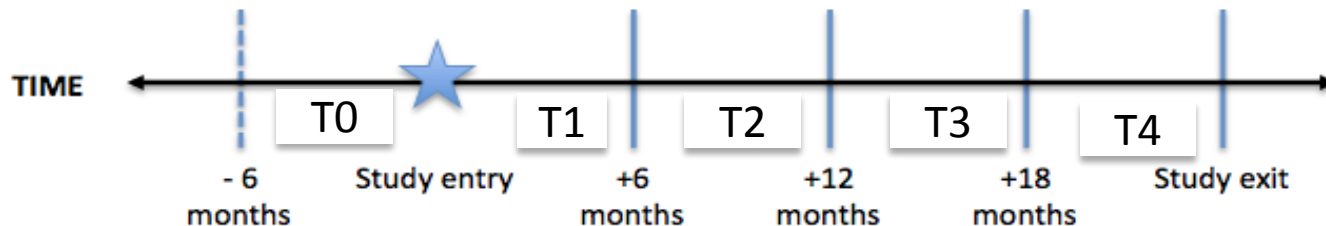
# Method

## Linear mixed effects models

- Random intercept for person
- Regression assumptions checked

## Modified intent-to-treat analysis

- baseline plus FB data from at least 1 other time point
- R package NLME, using RML



# Results: Hypothesis 1

**Table 2. Coefficients for the linear mixed models testing for change in % HAL between the treatment and control participants over time (N = 329)**

	Model 1			Model 2		
	Beta	CI	P	Beta	CI	P
Intercept	4.44	(3.51, 5.36)	0.00	0.15	(-2.50, 2.80)	0.91
T1	-0.69	(-1.87, 0.50)	0.26	-0.67	(-1.86, 0.51)	0.27
T2	-0.03	(-1.22, 1.17)	0.96	-0.01	(-1.21, 1.19)	0.99
T3	-0.42	(-1.61, 0.78)	0.49	-0.40	(-1.60, 0.79)	0.51
T4	0.40	(-0.8, 1.61)	0.51	0.41	(-0.79, 1.62)	0.50
Group treatment	-0.49	(-1.81, 0.82)	0.46	-0.45	(-1.76, 0.85)	0.50
T1*treatment	1.75	(0.06, 3.44)	<b>0.04*</b>	1.74	(0.05, 3.43)	<b>0.04*</b>
T2*treatment	0.95	(-0.75, 2.65)	0.27	0.94	(-0.76, 2.64)	0.28
T3*treatment	0.78	(-0.92, 2.48)	0.37	0.78	(-0.92, 2.48)	0.37
T4*treatment	-0.17	(-1.89, 1.56)	0.85	-0.18	(-1.91, 1.55)	0.84
Sex female				-0.17	(-1.04, 0.70)	0.70
Age				0.20	(0.09, 0.30)	<b>&gt;0.0001*</b>

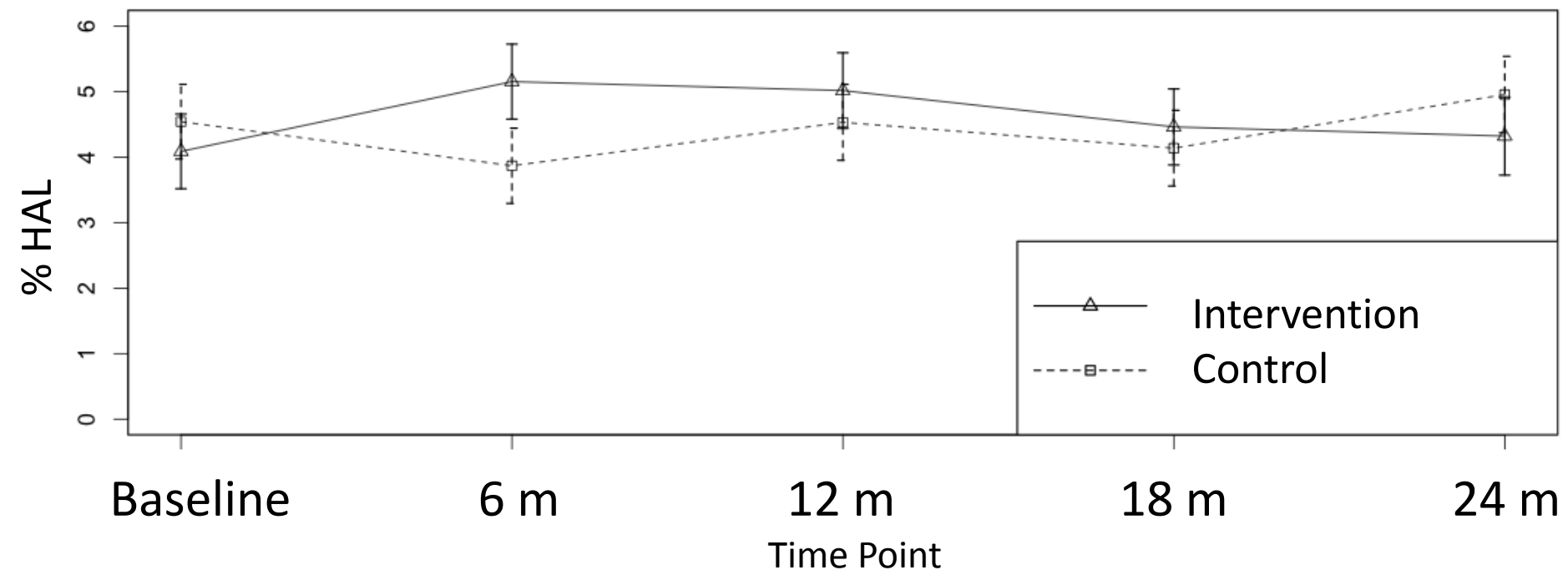
The reference categories are control (for group) and baseline/T0 (for time)

# Results: Hypothesis 1

4.85% of posts were classified as HAL

- $\text{Range}_{\text{HAL words/HAL post}} = 1 - 18$
- $M_{\text{HAL words/HAL post}} = 1.3 (\pm 0.8)$

Percent of Facebook posts about HAL by condition  
(Predicted means, standard errors)



# Results: Hypothesis 2

Engagement Dichotomized as:

- Minimally engaged:  $< 1/\text{mo}$
- Engaged:  $\geq 1/\text{mo}$

Highly variable engagement, decreasing over time

**Table 3. Engagement\* with the study's Facebook page over time<sup>^</sup>**

	T1	T2	T3	T4
Mean (SD)	18.09 (40.09)	11.35 (17.23)	14.30 (20.31)	8.66 (14.19)
Median	5	3	5	2
Range	1 – 285	1 – 84	1 – 92	1 – 68

# Results: Hypothesis 2

**Table 4. Coefficients for the linear mixed models testing for change in % HAL among treatment participants by study Facebook engagement status over time (N = 162)**

	Model 1			Model 2		
	Beta	CI	P	Beta	CI	P
Intercept	4.60	(3.62, 5.58)	0.00	-3.15	(-7.06,0.77)	0.12
T2	0.13	(-1.04, 1.29)	0.83	0.15	(-1.02,1.31)	0.80
T3	-0.03	(-1.21, 1.15)	0.96	0.00	(-1.18,1.18)	0.99
T4	-0.45	(-1.63, 0.72)	0.45	-0.45	(-1.62,0.73)	0.46
Group: engaged	1.56	(-0.25, 3.37)	<b>0.09</b>	1.54	(-0.26,3.34)	<b>0.09</b>
T2*engaged	-0.82	(-3.4, 1.75)	0.53	-0.86	(-3.43,1.71)	0.51
T3*engaged	-2.73	(-5.22, -0.25)	<b>0.03*</b>	-2.85	(-5.34,-0.37)	<b>0.03*</b>
T4*engaged	-1.13	(-4.20, 1.93)	0.47	-1.24	(-4.3,1.82)	0.43
Sex female				0.27	(-1.02,1.56)	0.68
Age				0.34	(0.18,0.49)	<b>&gt;0.0001*</b>

Minimally engaged (< 1 interaction on the study's Facebook page / month) is reference category for group

T1 (baseline) is reference category for time point

**When dichotomized, engagement on study's FB page not associated with posting about HAL as hypothesized**

# Results: Hypothesis 2

**Table. Coefficients for the linear mixed models testing for change in % HAL among treatment participants by study Facebook engagement score over time (N = 162)**

	Beta	CI	P
Intercept	-2.45	(-6.44, 1.54)	0.23
Time	-0.24	(-0.57, 0.09)	0.16
FB score	0.03	(0.01, 0.05)	<b>0.01*</b>
Sex female	0.28	(-1.01, 1.56)	0.67
Age	0.33	(0.18, 0.49)	0.00

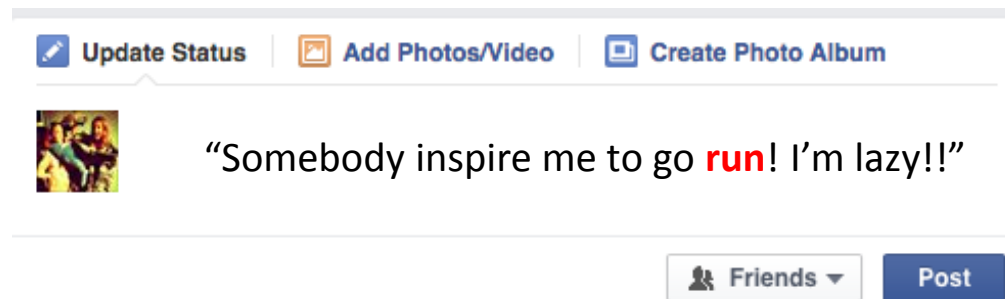
T1 (baseline) is reference category for time point

**Using a continuous FB engagement score, there is a small positive association between interacting on the study's FB page and posting about HAL**



# Discussion

- Treatment group shared more HAL content, but the effect did not persist over time
- There is limited support that the effect was explained by observable engagement on the study's FB page



# Discussion

## Strengths

- Communication with existing friends
- Iteratively derived and tested dictionary
- Examined change over time

## Limitations

- Limited dictionary power
- Dictionary only includes health enhancing behavior
- Engagement defined as observable engagement

## Lurking on Facebook

*“Just because I’m ‘passive’ doesn’t mean I’m ignoring it.”*

# Future Work

- Improve dictionary classifier
- Unsupervised classification approach
  - Topic model
  - Machine learning
- Look at diet and exercise separately
- Network effects
  - 214 friendships between study participants
  - 40% friendships between intervention and control participants

Extra Slides

# HAL Dictionary creation

## **Inclusion criteria**

- Purposeful exercise words
- Healthy food words

## **Exclusion criteria**

- Bigrams and beyond
- Hyphenated words
- Household chores
- Leisure activities (e.g., bowling, skydiving)
- Diet: spices

# HAL Dictionary examples

## Exercise

- **Activity descriptor:** aerobic
- **Activity:** runing
- **Activity tool:** bicycle
- **Races/competitions:** 5K

## Diet

- **Fruit:** banana
- **Veggie:** carrot
- **Food descriptor:** organic
- **Protein:** almond
- **Grain:** quinoa
- **Nutrient:** antioxidant



# Exercise words (N = 165)

10k	cheerleading	horsebackriding	raced	soulcycle	workouts
5k	climb	intramural	races	spartan	wrestling
abs	climbing	jazzercise	racing	spin	yoga
active	coach	jetski	racquetball	squash	zumba
aerobic	cricket	jetskiing	racquetballs	squats	
anaerobic	crossfit	jog	rafting	stairmaster	
backpacking	crunches	jogging	ran	stairs	
badminton	duathlon	karate	recipes	surf	
ballet	elliptical	kayaking	reps	surfboard	
baseball	endurance	kettlebell	rockclimbing	surfboards	
baseballs	ergometer	kettlebells	rollerblading	surfed	
basketball	exercise	kickball	rowing	surfing	
basketballs	exercises	kitesurf	rugby	swam	
biathlon	fencing	kitesurfing	run	swim	
bicep	fitness	lacrosse	runner	swimming	
bicycle	football	lbs	running	taekwando	
bicycled	footballs	lunge	situp	toughmudder	
bicycles	frisbee	lunges	situps	trainer	
bicycling	frisbees	mudder	skateboarding	training	
bike	golf	paddleball	skating	treadmill	
biked	gym	paddleboarding	ski	triathlon	
bikes	gymnasium	pedometer	skied	ups	
biking	gymnasiums	pilates	skiing	volleyball	
boadyboarded	gymnastics	plank	skijump	walk	
boadyboarding	gyms	planks	skijumping	walked	
bodyboard	hackysac	plyometric	skis	walking	
bodyboards	hackysacs	plyos	snowshoe	waterpolo	
bootcamp	handball	pullup	snowshoeing	weights	
bootcamps	healthy	pullups	snowshoes	windsurf	
boxing	hike	pushup	soccer	windsurfed	
cardio	hiking	pushups	softball	windsurfing	
cardiovascular	hockey	race	softballs	workout	

# HAL Diet (N = 186)

almond	cherries	lentils	persimmons	tilapia
almonds	cherry	lettuce	pineapple	tofu
antioxidant	chickpea	lowfat	pineapples	tomato
antioxidants	chickpeas	mango	pistachio	tomatoes
apple	coconut	mangoes	pistachios	trailmix
apples	coconuts	mangos	plate	tuna
apricot	cod	melon	plum	vegan
apricots	corn	melons	plums	vegetable
artichoke	cranberries	muesli	pomegranate	vegetables
artichokes	cranberry	mushroom	pomegranates	vegetarian
arugula	cucumber	mushrooms	potassium	veggie
asparagus	cucumbers	nectarine	potato	veggies
avocado	currants	nectarines	potatoes	vitamin
avocados	eggplant	nonfat	protein	vitamins
banana	eggplants	nutrient	prune	walnut
bananas	fiber	nutrients	prunes	walnuts
barley	fig	oat	quinoa	water
bean	figs	oatmeal	radish	watercress
beans	fish	oats	radishes	watermelon
beet	flax	okra	raisin	watermelons
beets	fruit	olive	raisins	wellness
blackberries	fruits	olives	rasberries	wholewheat
blackberry	grain	onion	rasberry	yogurt
blueberries	grains	onions	recipe	zucchini
blueberry	granola	orange	recipes	
bran	grape	oranges	romaine	
broccoli	grapefruit	organic	salad	
cabbage	grapefruits	papaya	salads	
cabbages	grapes	papayas	salmon	
calcium	guava	parsnip	soy	
cantaloupe	health	pasta	soybean	
cantaloupes	healthy	peach	spinach	
carrot	honeydew	peaches	sprout	
carrots	iron	peanut	sprouts	
cashew	kale	peanuts	squash	
cashews	kiwi	pear	strawberries	
cauliflower	kiwifruit	pears	strawberry	
celery	leek	pecan	swordfish	
cereal	leeks	pecans	tangerine	
chard	lentil	persimmon	tangerines	

# Human Coding

## 40 disagreements (out of 2,614):

- Context Unclear: 16  
→ conservative vs. generous coding
- Human Error: 11
- Vague Plan: 6
- Questionable Purpose: 5
- Request for Support: 4

## Words Used out of context:

**Nutrition:** apple, banana, nutrition, soy (Spanish)

**Exercise:** baseball, basketball, football, dancing

# Recruitment

## Inclusion Criteria

- Aged 18 to 35 years
- BMI  $\geq 25$  and  $\leq 34.9$  kg/m<sup>2</sup>
- Owned a personal computer
- Owned a mobile phone and used text messaging
- Facebook user or willing to begin

## Exclusion Criteria

- Clinically diagnosed comorbidities
- Psychiatric or medical conditions
- Prescribed dietary or physical activity changes
- Taking medications that altered weight
- Pregnant or intending to be within two years

UC San Diego



SAN DIEGO STATE  
UNIVERSITY



California State University  
SAN MARCOS

# Results: Hypothesis 1

## **Males talked more about exercise**

- 6 months into the study, males in the treatment group post significantly more exercise HAL than females in the control group
  - (Beta = - 3.68; SE = 1.72;  $p < 0.05$ )

## **Females talked more about diet**

- 12 months into the study, females in the treatment group post significantly more diet HAL than males in the control group
  - (Beta = 2.41; SE = 0.96;  $p < 0.05$ )

# Weight Loss

- **DV:** weight (kg)
- **IV:**
  - % of posts that were HAL
  - % of social support that was for HAL posts
- **Covariates:** age, sex, group assignment
- **Time:** 2 years
  - T1 (baseline); T2 (6 months)... T5 (24 months)
- **Linear mixed effects models:**
  - Random intercept for person
  - Regression assumptions checked and met
- **Analysis:** R package NLME, using RML



# Weight Loss

**Table 2. Coefficients for the linear mixed models testing for association between receiving social support for talking about HAL on Facebook and change in weight (kg) over time**

	Model 1			Model 2			Model 3		
	Beta (CI)		P	Beta (CI)		P	Beta (CI)		P
Intercept	80.40	(78.96, 81.84)	0.00	80.43	(78.99, 81.87)	0.00	82.98	(75.44, 90.52)	0.00
T2	0.08	(-0.65, 0.80)	0.84	0.08	(-0.64, 0.80)	0.83	0.05	(-0.67, 0.77)	0.89
T3	-0.25	(-0.96, 0.45)	0.48	-0.23	(-0.94, 0.48)	0.53	-0.25	(-0.96, 0.46)	0.49
T4	0.49	(-0.24, 1.22)	0.19	0.50	(-0.23, 1.23)	0.18	0.48	(-0.25, 1.21)	0.20
T5	0.96	(0.23, 1.69)	0.01	0.97	(0.24, 1.70)	0.01	0.94	(0.21, 1.66)	0.01
% HAL SS	0.06	(-0.02, 0.14)	0.13	0.07	(-0.01, 0.16)	0.09	0.07	(-0.01, 0.16)	0.10
T2 * % HAL <sub>ss</sub>	-0.08	(-0.18, 0.02)	0.11	-0.08	(-0.18, 0.02)	0.11	-0.08	(-0.18, 0.02)	0.13
T3 * % HAL <sub>ss</sub>	-0.04	(-0.13, 0.05)	0.38	-0.04	(-0.13, 0.05)	0.34	-0.04	(-0.13, 0.05)	0.37
T4 * % HAL <sub>ss</sub>	-0.09	(-0.18, 0.01)	<b>0.07</b>	-0.09	(-0.19, 0.01)	<b>0.06</b>	-0.09	(-0.18, 0.01)	<b>0.07</b>
T5 * % HAL <sub>ss</sub>	-0.04	(-0.13, 0.05)	0.41	-0.04	(-0.13, 0.05)	0.38	-0.04	(-0.13, 0.06)	0.44
% HAL				-0.03	(-0.09, 0.04)	0.44	-0.03	(-0.10, 0.03)	0.36
Sex female							-14.88	(-17.43, -12.34)	0.00
Age							0.38	(0.07, 0.68)	0.02
Group treatment							-0.77	(-3.06, 1.51)	0.51

The reference categories is baseline/T1 (for time)

**Receiving social support on HAL posts not associated with weight loss.**

# Weight Loss

**Table 3. Coefficients for the linear mixed model testing for association between receiving social support for talking about HAL on Facebook and change in weight (kg) by sex in the treatment group**

	Female treatment group			Male treatment group		
	Beta (CI)		P	Beta (CI)		P
Intercept	68.14	(53.92, 82.35)	0.00	86.31	(70.82, 101.79)	0.00
T2	0.12	(-1.02, 1.26)	0.84	-1.98	(-3.97, 0.00)	0.05
T3	-0.17	(-1.31, 0.96)	0.76	-2.88	(-4.86, -0.89)	0.01
T4	0.59	(-0.56, 1.74)	0.31	-0.87	(-3.17, 1.43)	0.46
T5	1.35	(0.13, 2.57)	0.03	-0.44	(-2.41, 1.54)	0.66
% HAL SS	0.16	(0.00, 0.32)	0.05	0.04	(-0.15, 0.23)	0.70
T2 * % HALss	-0.20	(-0.37, -0.04)	<b>0.02*</b>	0.01	(-0.20, 0.22)	0.92
T3 * % HALss	-0.13	(-0.28, 0.03)	0.12	0.07	(-0.14, 0.28)	0.51
T4 * % HALss	-0.13	(-0.30, 0.03)	0.11	-0.08	(-0.33, 0.17)	0.55
T5 * % HALss	-0.11	(-0.28, 0.07)	0.22	-0.03	(-0.22, 0.15)	0.72
% HAL	-0.07	(-0.21, 0.08)	0.36	-0.06	(-0.22, 0.11)	0.50
Age	0.36	(-0.28, 1.00)	0.27	0.22	(-0.42, 0.85)	0.51

The reference categories is baseline/T1 (for time)

**For every 20% increase in social support on HAL posts, females in the treatment group lost 9 lbs from baseline to 6 months. Effect did not persist.**