

The Persuasive Power of Human-Machine Dialogue

Divya Ramachandran
John Canny

University of California, Berkeley

BiD

BERKELEY INSTITUTE OF DESIGN

Persuasive 2008
Oulu, Finland

June 6, 2008



Alice

Mom, my friend Betsy is going to a rock concert next Saturday and invited me to go along. It's a concert by one of my favorite bands and I really want to go. It's going to be in the city about an hour away. It'll end pretty late at night so it would be best if one of us drove. Betsy's parents are taking her car on a trip, so I'll need to borrow your car.

Goal: To borrow mom's car to attend a rock concert.

Motivation

BiD



Alice

Mom, my friend Betsy is going to a rock concert next Saturday and invited me to go along.

Where is it?

It's going to be in the city about an hour away

How are you going to get there?

It'll end pretty late at night so it would be best if one of us drove.

Is Betsy driving?

Betsy's parents are taking her car on a trip, so I'll need to borrow your car.

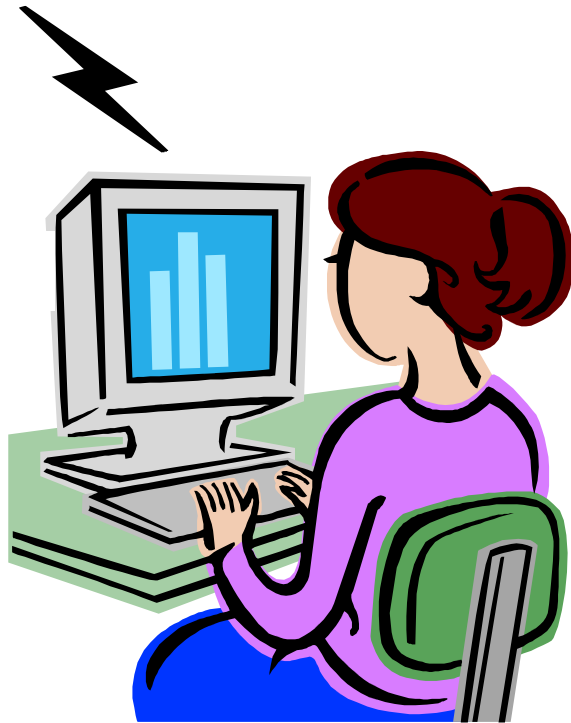
Which method should Alice choose?

- (1) All the information right up front
- (2) Short relevant responses to prompts

The Question (Version 1.0)



How does this translate to the design of persuasive interfaces?



Recorded Message System



Dialogic System

Study Overview



1

52 subjects, university students and staff
Compared persuasive power of two speech-based systems

2

Used NASA Moon Survival task for measuring persuasion
Participants rank 15 items in order of priority for survival on moon

3

Participants accessed information using one of two systems, which served as a confederate in the experiment

4

After interacting with system, participants ranked 15 items again

5

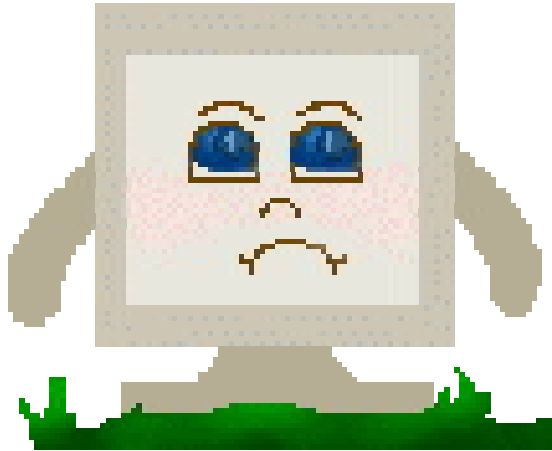
Compared persuasive power of both systems by analyzing change in rankings

6

RESULTS!

Related Work

BiD



gender
personality
accent
emotion



likeability
trustworthiness
informativeness
competence
credibility

[Nass et al. '05]

Tailoring Personal Information

[Dijkstra '06], [Davis, '07], [Kraft et al. '07], [Brug et al. '96]



Recruitment

Berkeley's eXperimental Social Science Laboratory

Subjects

52 total subjects

40% male, 60% female

46 university students (21 yrs), 6 staff members (33 yrs)

Conditions

Recorded Message System

24 subjects

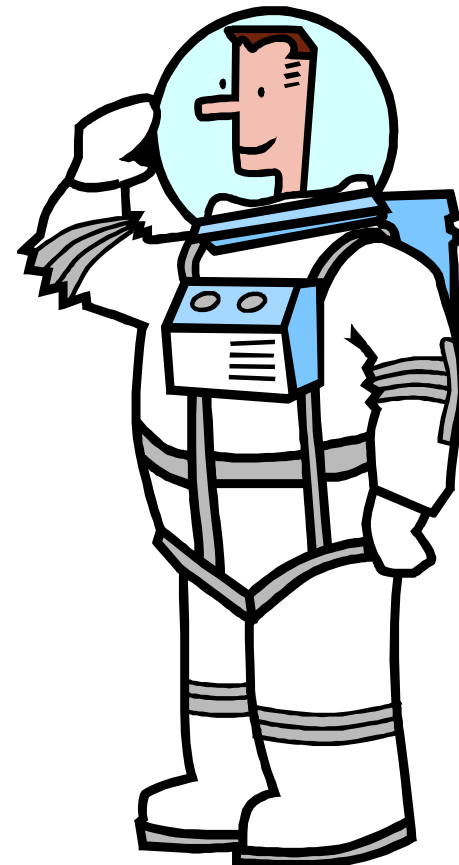
Dialogic System

28 subjects

NASA Moon Survival Problem

BiD

5 gallons of water
Food concentrate
Stellar map (of moon's constellations)
Box of matches
50 ft of nylon rope
First aid kit with injection needles
Life raft
Portable heating unit
2 hundred-pound tanks of oxygen
Signal flares
Two .45 caliber pistols
Parachute silk
1 case dehydrated pet milk
Magnetic compass
Solar-powered FM receiver-transmitter



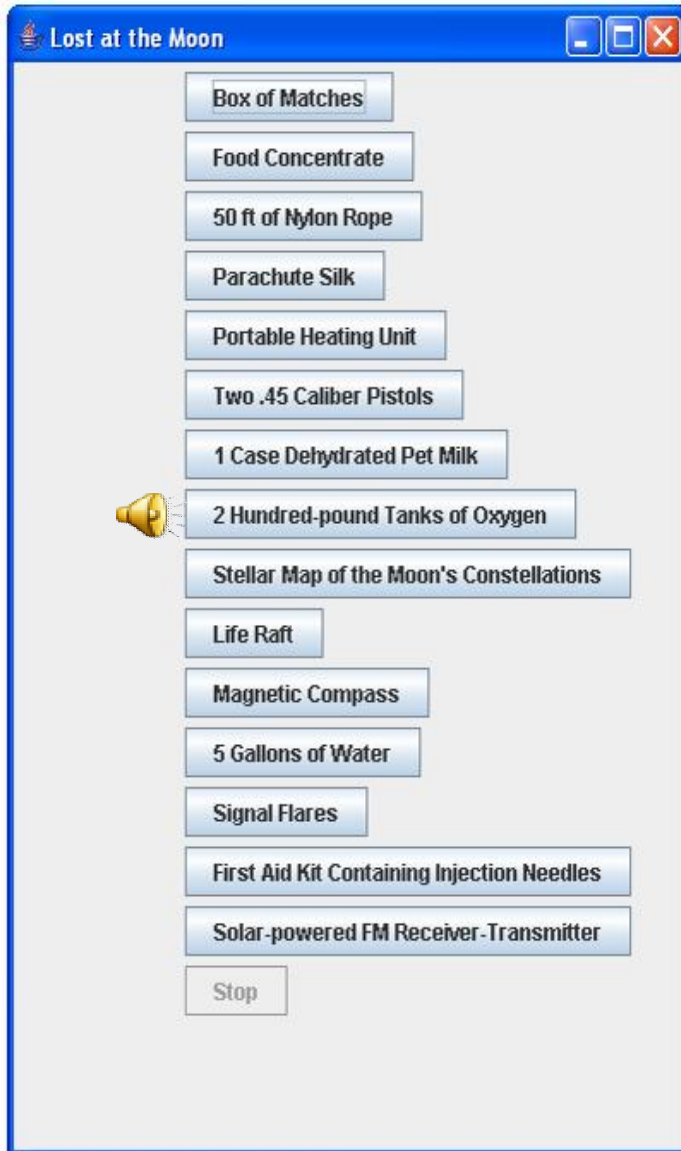


Recorded Message System



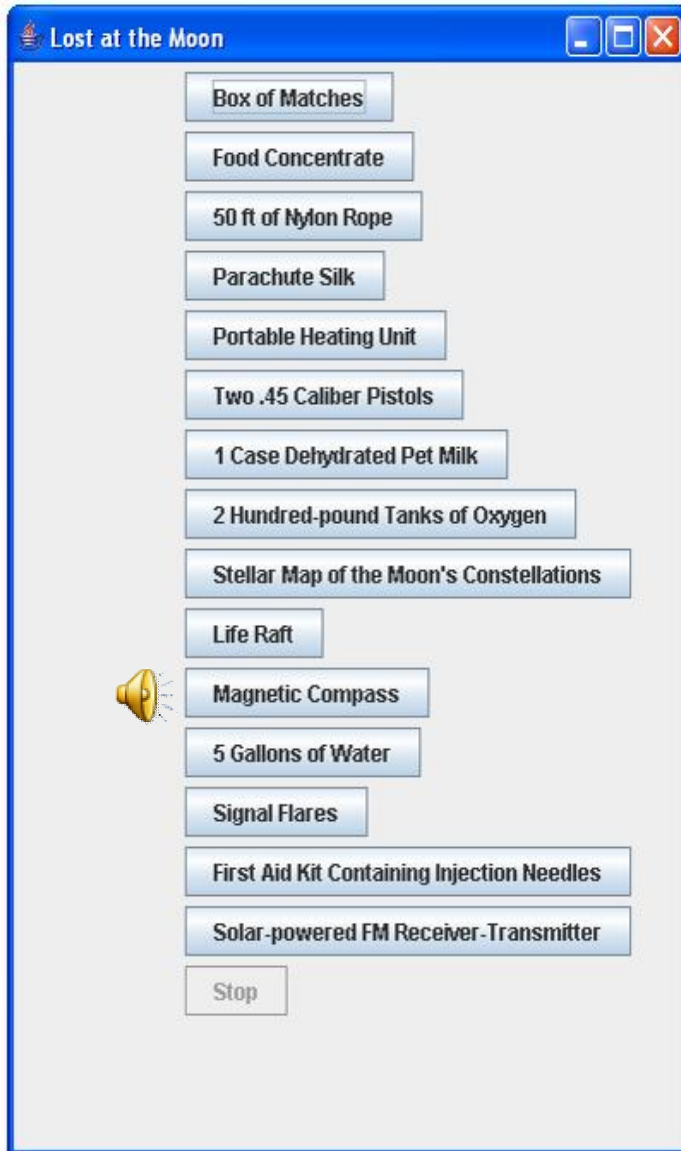
Dialogic System

System A: Recorded Message



This is clearly the most important item on the list. Oxygen is absolutely essential for life support. There's no way we'll survive the trek to the mother ship without it, even if we know where it is. We'll die without it. Oxygen is number one.

System A: Recorded Message



It's so crucial to find our way back to the mother ship before our resources like oxygen run out. We need everything possible to help us find our way to the mother ship. Clearly the compass is the only tool we've got to point us in the right direction. So the order should go signal flares for the mother ship to find us and then magnetic compass for us to at least head in the right direction towards it.

System B: Dialogic

BiD

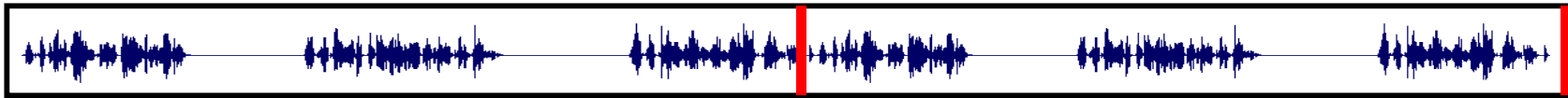


How could we use the signal flares? 📢

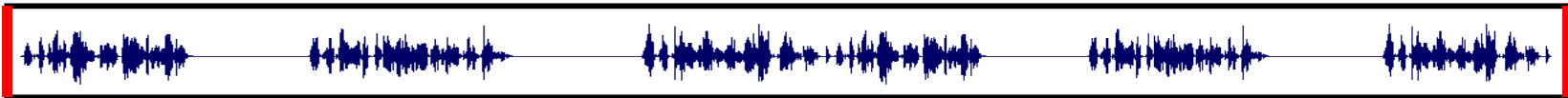
Will the life raft be useful? 📢

Wizard-of-Oz System Design

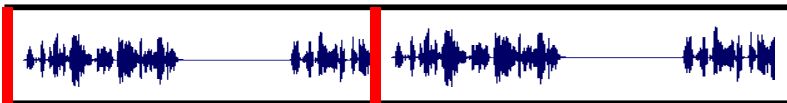
BiD



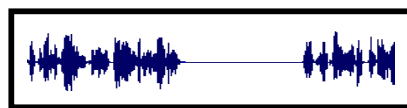
This is clearly the most important item on the list. Oxygen is absolutely essential for life support.



There's no way we'll survive the trek to the mother ship without it, even if we know where it is.



We'll die without it. Oxygen is number one.



We'll die without it.



- I. The dialogic system will be more persuasive than the recorded message system.

- II. Participants using the dialogic system will receive less information than those using the recorded message system.

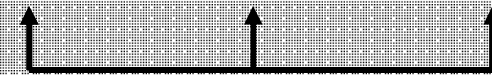
Expert Solution



Expert NASA Ranking

- 2 hundred-pound tanks of oxygen
- 5 gallons of water
- Stellar map (of moon's constellations)
- Food concentrate
- Solar-powered FM receiver-transmitter
- 50 ft of nylon rope
- First aid kit with injection needles
- Parachute silk
- Life raft
- Signal flares
- Two .45 caliber pistols
- 1 case dehydrated pet milk
- Portable heating unit
- Magnetic compass
- Box of matches

E	Pre	Post
1	1	1
2	3	3
3	4	2
4	7	4
5	11	12
6	2	14
7	5	11
8	6	8
9	15	5
10	10	7
11	13	10
12	9	6
13	8	9
14	14	13
15	12	15



$$c(E, \text{pre}) \times c(E, \text{post})$$

Standard Rank Correlation Metrics

- Spearman Rank Coefficient
 - Kendall's Tau
 - Spearman's Footrule
- } Asymptotically normal as n (# of items) grows
- } Better for smaller n

Individual limitations suggest use of multiple methods for rank analysis [Diaconis '77]

$$\textit{persuasive power} = c(E, \textit{pre}) - c(E, \textit{post})$$

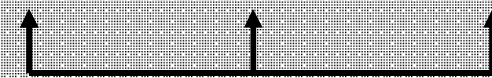
Confederate System Ranking



Confederate System Ranking

2 hundred-pound tanks of oxygen
Signal flares
Magnetic compass
Food concentrate
Solar-powered FM receiver-transmitter
Two .45 caliber pistols
50 ft of nylon rope
First aid kit with injection needles
Stellar map (of moon's constellations)
1 case dehydrated pet milk
5 gallons of water
Portable heating unit
Parachute silk
Life raft
Box of matches

C	Pre	Post
1	1	1
2	3	3
3	4	2
4	7	4
5	11	12
6	2	14
7	5	11
8	6	8
9	15	5
10	10	7
11	13	10
12	9	6
13	8	9
14	14	13
15	12	15



$c(C, \text{pre}) \setminus (C, \text{post})$

A New Metric



Incorporate movement of user i away from expert *and* towards confederate

$$DL_i = \sum_j (C_j - E_j)(post_{ij} - pre_{ij})$$

This gives rank changes $(post_{ij} - pre_{ij})$ unequal weights $(C_j - E_j)$

$$DS_i = \sum_j sign(C_j - E_j)(post_{ij} - pre_{ij})$$

Rank Metrics Used

Spearman Rank Coefficient

Kendall's Tau

Spearman's Footrule

DS

Comparison of Persuasive Force across Systems

Measure	Effect Size*	P-value (t-test)	P-value (Perm)
Spearman Rank Coefficient	0.5011	$p = 0.0379$	$p = 0.0380$
Kendall's Tau	0.4612	$p = 0.0507$	$p = 0.0527$
Spearman's Footrule	0.5849	$p = 0.0193$	$p = 0.0191$
DS	0.5977	$p = 0.0176$	$p = 0.0181$

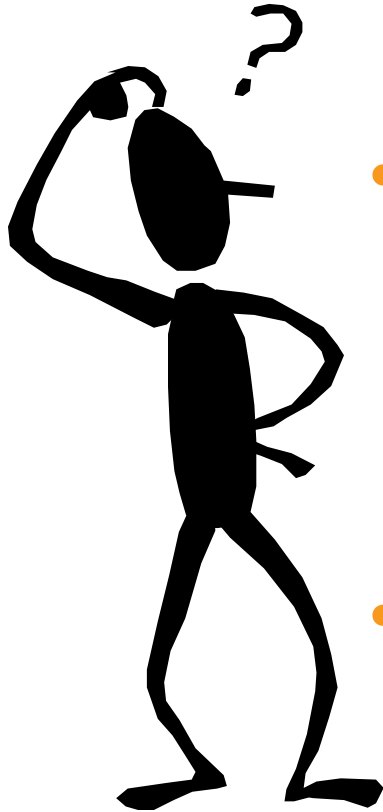
*Effect size = difference between persuasive forces ($\mu_{\text{dialogic}} - \mu_{\text{recorded message}}$)

Comparison of Amount of Information Heard across Systems

Number of Statements	Effect Size	P-value (t-test)
102	3.223	$p < 10e^{-17}$

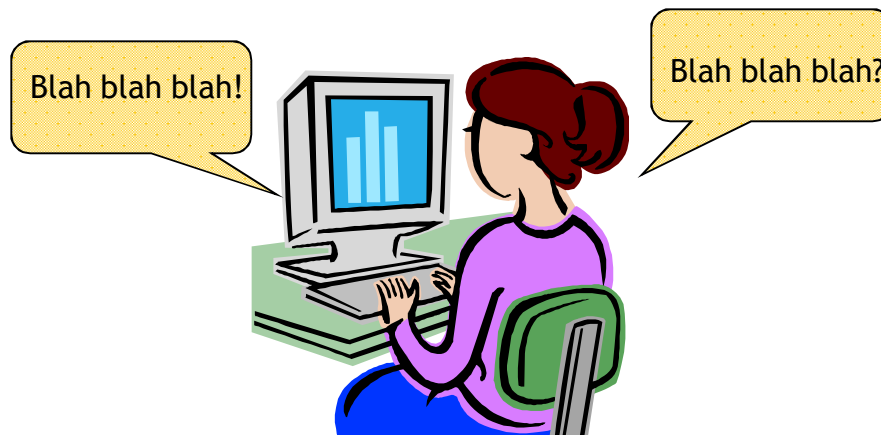
I. The dialogic system will be more persuasive than the recorded message system.

II. Participants using the dialogic system will receive less information than those using the recorded message system.



- Wizard-of-oz system
 - subjects could have suspected the technique
 - experimenter bias
 - better recognition than the state-of-the-art
- *DS* measure needs further cross-validation

- Dialogic style of information presentation is more persuasive than lecture style
- Dialogic interfaces could add to the effectiveness of persuasive technologies
- The *DS* measure may be a more sensitive method of interpreting user rankings of few items



Acknowledgements



- eXperimental Social Science Laboratory staff
- Colleagues at Berkeley Institute of Design, especially Dave Nguyen for help in study design
- Our reviewers for their valuable feedback
- Lisa Rowland (the voice)
- Our participants

Thank you!

Speech & Persuasion Experiment

[Ramachandran '08]

BiD

4

After interacting with system, ranked 15 items again

5

An “expert” ranking of items has been published by NASA

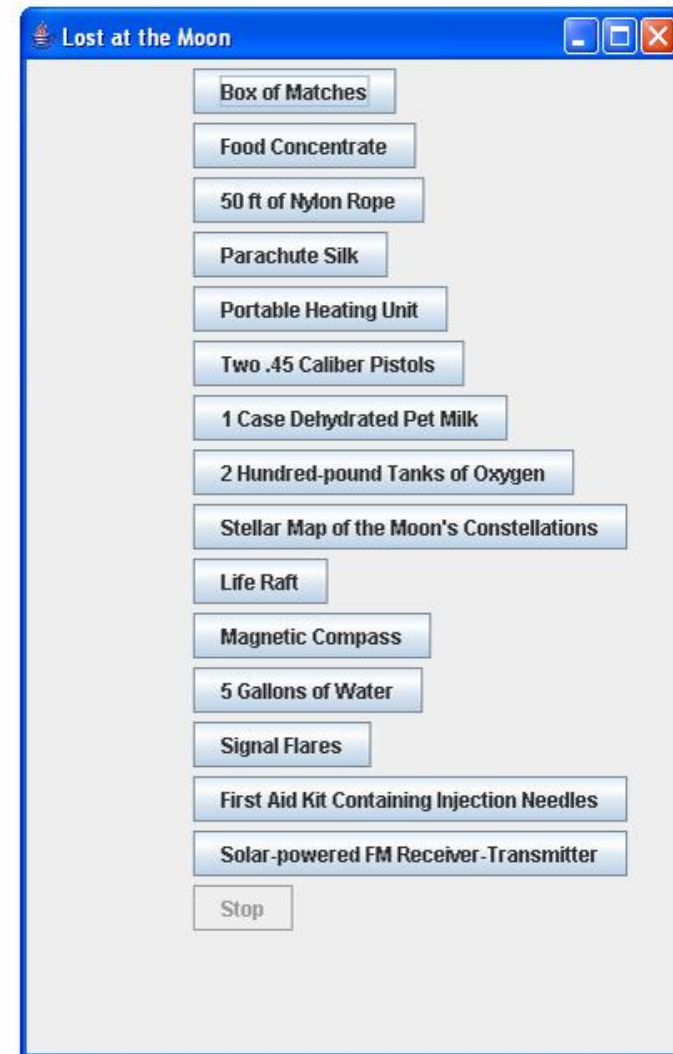
Distance moved away from “expert” solution used as measure of persuasion

6

Compared persuasiveness of both systems
Dialogic more persuasive with ($p < 0.05$)



System A: Dialogic



System B: Recorded Messages

Speech & Persuasion Experiment [Ramachandran 08]



4

After interacting with system, ranked 15 items again

5

An “expert” ranking of items has been published by NASA

Distance moved away from “expert” solution used as measure of persuasion

6

Compared persuasiveness of both systems
Dialogic more persuasive with ($p < 0.05$)



System A: Dialogic



System B: Recorded Messages