Long-Term Human-Computer Relationships

Timothy Bickmore
Assistant Professor
College of Computer and Information Science
Northeastern University, Boston

Overview

- Relational agents
- Lessons learned
- Automated health conscience

Relational Agents

- Computational artifacts designed to establish long-term, social-emotional relationships with users.
- Possible Components
  - Relational model
  - Relationship assessment
  - Relational behaviors
  - Persistent memory
  - User Identification

3

Why?

- Therapeutic/Working alliance has significant effects on
  - Patient satisfaction
  - Patient adherence
  - Outcomes

Relevance to ILLLC

- K-6 peer learning is better among friends than strangers
  - Friends engage in more extensive discourse, offer suggestions more readily, are more supportive and more critical than non-friends, work longer on task.
- Teacher empathy and caring (6th-8th graders)
  - Significant predictor of student motivation, pro-social goals, academic effort, GPA.
- In any longitudinal intervention, adherence to the intervention regimen is a pre-requisite for obtaining outcomes.

Human Relational Behavior

- Social Psychology
  - Social penetration theory / self-disclosure
  - Meta-relational communication
  - Continuity behaviors
- Helping & Psychotherapy
  - Unconditional positive regard
  - Empathic listening

- Sociolinguistics
  - Politeness theory
- Linguistics / Conversation Analysis
  - Structure & function of social dialogue
- Communication
  - Comforting behavior
  - Nonverbal immediacy behavior
- Change Over Time
  - Increasing common ground
  - Increasing intimacy
  - Decreasing politeness

6

Platform: Embodied Conversational Agents

- Emulate human face-to-face conversation
- Use nonverbal communicative behavior
  - gaze, posture, gesture, etc.

7

Demo: Virtual Hospital Discharge Nurse

- In-hospital patient education about post-discharge self-care.
- Primarily at time of discharge, but also have contact every day patient is in hospital.
- Research Foci:
  - Relational behavior ("bedside manner").
  - Explanation of medical documents.

Some lessons learned...

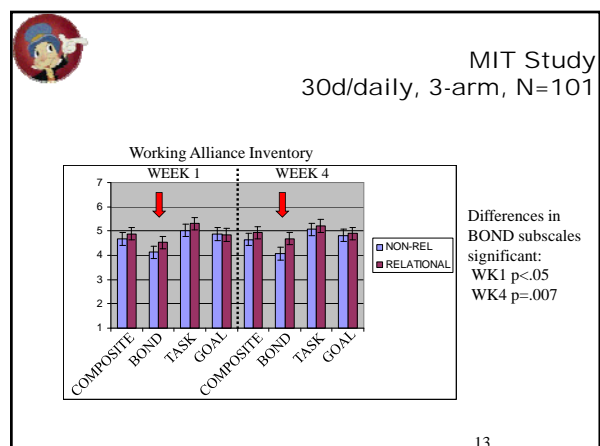## Lesson 1: The kitchen sink approach seems to work

Kitchen Sink Approach Works

- MIT, BMC studies demonstrated that use of
  - Social dialogue
  - Self disclosure
  - Meta-relational dialogue
  - Increasing common ground
  - Empathy
  - Nonverbal immediacy behavior
  - Humor
  - etc. etc.

  has significant impact on some self-reported and objective measures of relationship.

12

MIT Study
30d/daily, 3-arm, N=101

Working Alliance Inventory

WEEK 1   WEEK 4

COMPOSITE BOND TASK GOAL   COMPOSITE BOND TASK GOAL

NON-REL
RELATIONAL

Differences in BOND subscales significant:
WK1 p<.05
WK4 p=.007

13

Relational Results

- Self report:
  - "How much do you like Laura?" p<.05
  - "How would you characterize your relationship with Laura?" p<.05
  - "Would you like to continue working with Laura?" (at end of WK4) p<.05
- Behavioral measure:
  - Sentimental farewell (70% REL vs. 32% NON-REL; p<.001)

MIT Study

Educational Content pages viewed per session

PAGES/SESSION

Significant difference in educational pages viewed:
CONTROL < AGENT
p<.05

15

BMC Geriatrics Study

BMC GAP Exercise Study
21 participants
Age 63-85 (mean 74)
76% African American
86% Low Reading Literacy
38% Never Used a Computer
29% Used Computer a "Few Times"

16

Geriatrics Results: Ease of Use

Easy — 1.9 — Difficult
1 2 3 4 5 6 7

"That is so easy. That is so good. Regular computers I don't do. But, that was so easy, even a baby could do that."

Results: Satisfaction

Not at all — 6.3 — Very
1 2 3 4 5 6 7
Satisfaction with Overall System

Not at all — 5.4 — Very
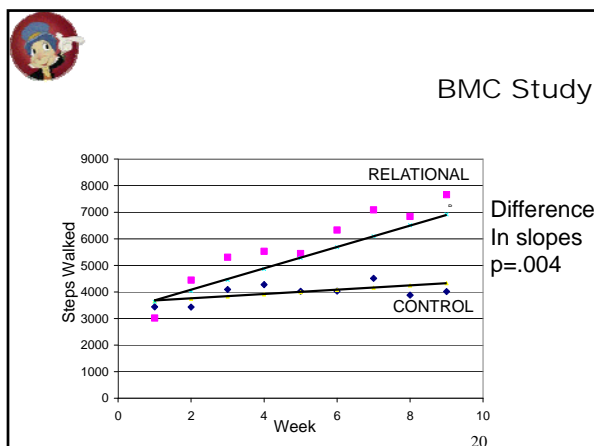1 2 3 4 5 6 7
Satisfaction with Laura

Not at all — 6.4 — Very
1 2 3 4 5 6 7
Desire to Continue
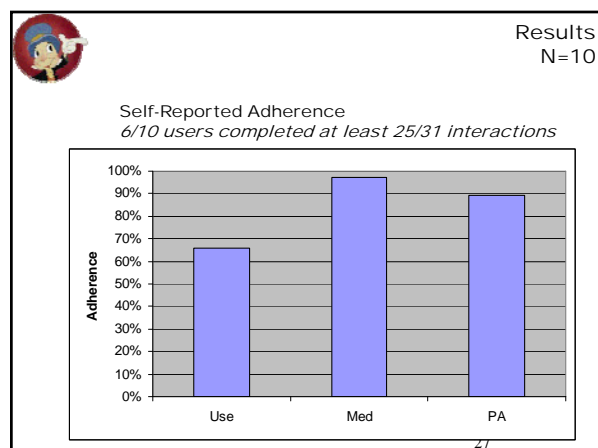
18

Results: Relationship with Laura

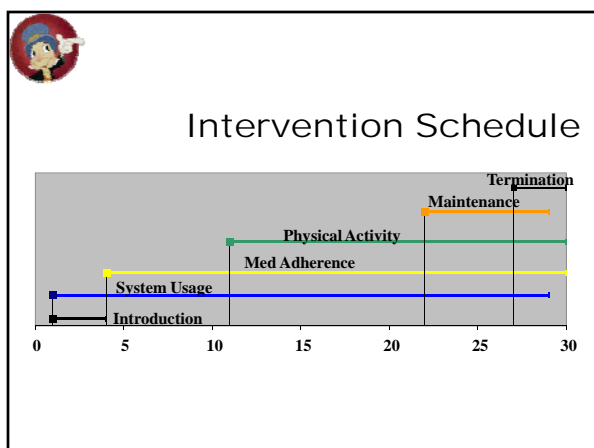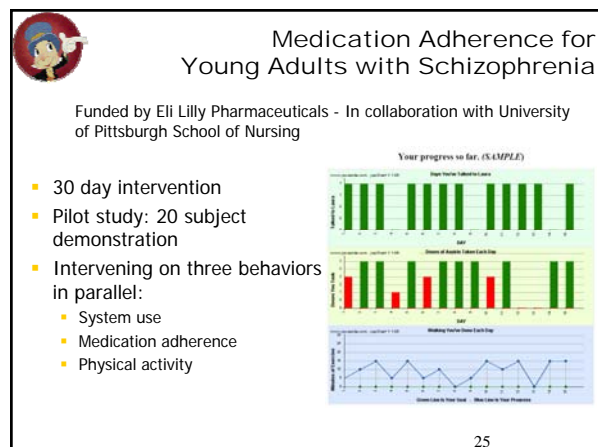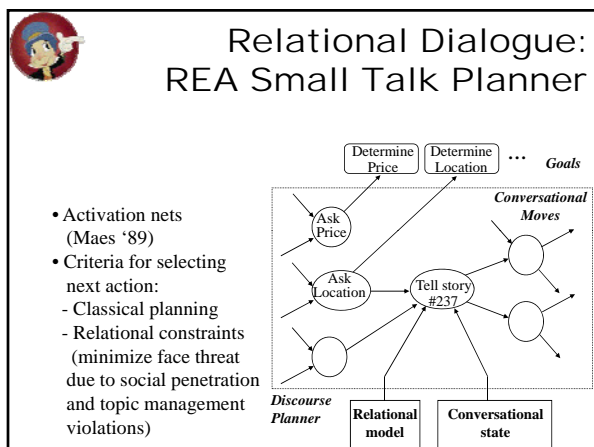Stranger — 5.6 — Close Friend
1 2 3 4 5 6 7
Relationship with Laura

Not at all — 6.4 — Very
1 2 3 4 5 6 7
Trust in Laura

Not at all — 6.3 — Very
1 2 3 4 5 6 7
Liking of Laura

19

BMC Study

Steps Walked vs Week. RELATIONAL, CONTROL. Difference In slopes p=.004

20

Lesson 2.
Relationship can be treated as a behavioral variable

Relational Dialogue:
REA Small Talk Planner

- Activation nets (Maes '89)
- Criteria for selecting next action:
  - Classical planning
  - Relational constraints (minimize face threat due to social penetration and topic management violations)

Determine Price | Determine Location | ... *Goals*

Conversational Moves

Ask Price

Ask Location → Tell story #237

Discourse Planner

Relational model | Conversational state

Medication Adherence for
Young Adults with Schizophrenia

Funded by Eli Lilly Pharmaceuticals - In collaboration with University of Pittsburgh School of Nursing

Your progress so far. *(SAMPLE)*

- 30 day intervention
- Pilot study: 20 subject demonstration
- Intervening on three behaviors in parallel:
  - System use
  - Medication adherence
  - Physical activity

25

Intervention Schedule

Termination
Maintenance
Physical Activity
Med Adherence
System Usage
Introduction

0   5   10   15   20   25   30

Results
N=10

Self-Reported Adherence
6/10 users completed at least 25/31 interactions

Use | Med | PA

27

4

Results

Attitudes

Bar chart with categories: Satisfaction, Trust, Relationship, Easy, Continue. Y-axis from 0.00 to 5.00.

Lesson 3
Variability is important, repetitiveness is bad

MIT & BMC Studies
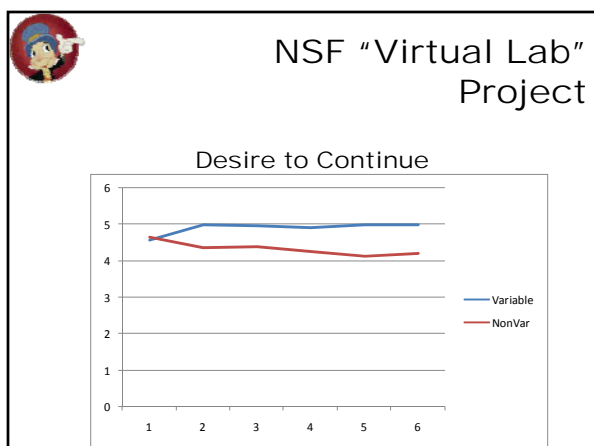
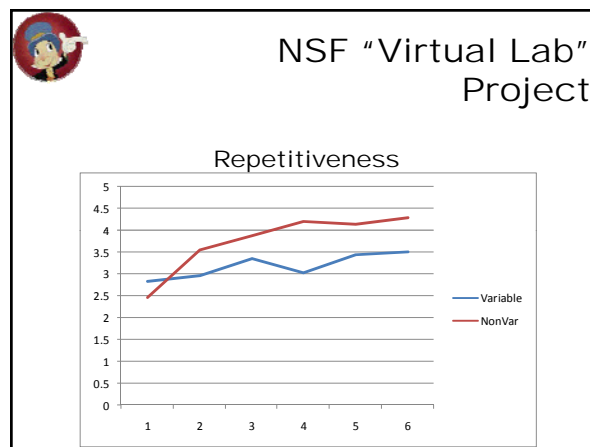- Repetitiveness: Most frequent complaint

The first couple of days I was impressed by it. But, there didn't seem to be a lot of variety going on after that, so it kind of lost my interest, it lost the engagement factor. Maybe, six or seven days into the study I could almost predict what she was going to say, and once the engagement was lost you sort of lose the power of the animated instructor.

Like 15 days into the study when I could almost predict what she was going to say, it became easier to do things like check my mail in between her responses. … Even with just little bits of variety your mind doesn't shut off.

In the beginning I was extremely motivated to do whatever Laura asked of me, because I thought that every response was a new response. Whereas, towards the end I could tell what she was going to say to a couple of my responses.

It would be great if Laura could just change her clothes sometimes.

30

NSF "Virtual Lab" Project

Repetitiveness

Line chart with Variable (blue) and NonVar (red) lines, x-axis 1–6, y-axis 0–5.

NSF "Virtual Lab" Project

Desire to Continue

Line chart with Variable (blue) and NonVar (red) lines, x-axis 1–6, y-axis 0–6.

NSF "Virtual Lab" Project

- More complaints from non-variable group.

I'm writing because the program "Karen" is getting very repetitive, and because of this kind of boring. For the past 5 days, "Karen" has been asking about the weather, and about walking outside. Now I know that the purpose of the program is to get people to walk, but it's almost like I want to click the answers quickly, so that I don't have to listen to her repetitiveness.

- 2 withdrawals from non-variable group
  - 0 from variable group

Lesson 4
Relational behavior doesn't make up for unreliability

PDA Context-awareness Study

- Compared automatic sensing of walking to explicit user signaling of walk start & end.
- Eight subject, 2-treatment (4day ea), within-subjects design.
- Results:
  - Awareness led to greater social bonding, but less walking.
  - Likely due to low perceived reliability
    - "There were times that I wasn't sure if it was picking up if I was walking briskly ..."
    - "I wasn't sure if it was always there."

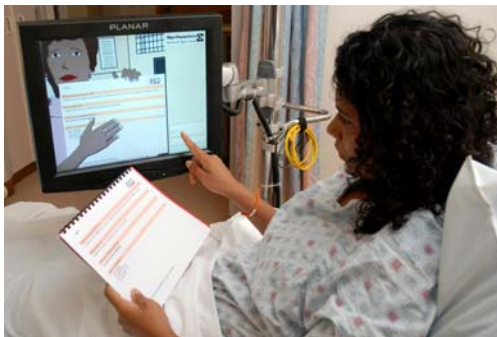### Tinker: Boston Museum of Science Tour Guide

- Biometric hand-print ID
  - Recognizes return visitors
  - Resumes dialogue & relational state
- Visitors really liked it when Tinker greeted them by name
- They hated it when she mis-recognized them

Lesson 5. Building relationships is easy when your users are starved for attention

---

Virtual Hospital Discharge Nurse

Relational Behavior aka "bedside manner"

- Randomized VN use of relational behavior
- Relational VN rated as
  - More caring, $t(17) = 1.9$, $p=.07$
  - Providing more useful information, $t(17) - 2.21$, $p<.05$

- *"She kept asking if I was tired, if I wanted to take a break. she cared about me, you know."*
- *"She treated me like a real person! She's not like a computer. This is awesome work! This is really excellent."*

Time for Caring

- Most appreciated the amount of information and time given to them by the agent.

- *"I prefer Louise, she's better than a doctor, she explains more, and doctors are always in a hurry."*
- *"It was just like a nurse, actually better, because sometimes a nurse just gives you the paper and says 'Here you go.' Elizabeth explains everything."*

"Who would you rather receive discharge instructions from?"

Doctor or Nurse 10%
Either 16%
Agent 74%
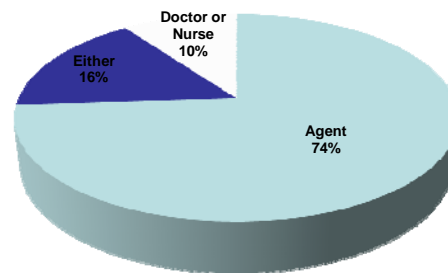
Imagine

- A Digital Conscience that goes everywhere with you
- Can sense when you are doing things that might have negative long-term consequences
- Whispers suggestions in your ear

Jiminy Cricket

- Earliest conceptualization of a "wearable conscience"?
  - Le avventure di Pinocchio, Carlo Collodi, 1881
- Jiminy did not fare well in original story:
  - Pinnochio strikes and kills the cricket with a hammer in their first meeting after he provides unwanted counsel.

PDA Platform

- Dell Axim X30
- Extended life battery
- EcerTech TiltControl 2D accel
- 1GB Flash drive
- Custom plastic case

"Just in Time Information for Exercise Adoption" Funded by NIH National Library of Medicine
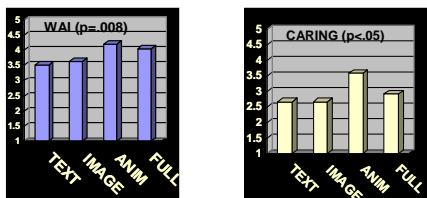
Modality Study

- Compared 4 modalities:
  - Text only
  - Text + Static agent image
  - Animated agent
  - Animated agent + nonverbal sounds
    - Backchannels, Discourse markers, etc.

Modality Study

WAI (p=.008)

TEXT | IMAGE | ANIM | FULL

CARING (p<.05)

TEXT | IMAGE | ANIM | FULL

- Animated agent also scored higher (approaching significance) on *credibility of health information* and *comfort using in the workplace.*
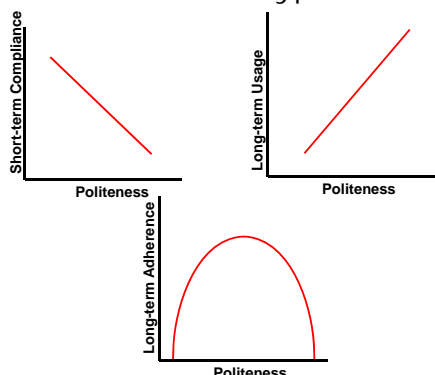
Interruption Studies

- What is the best way to interrupt people at work in order to motivate them to perform a healthy behavior?
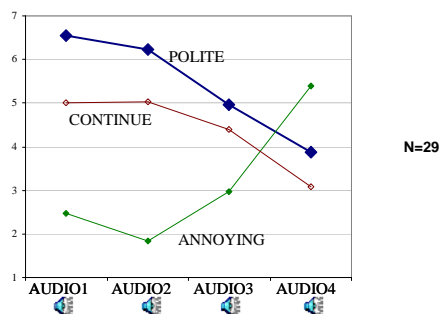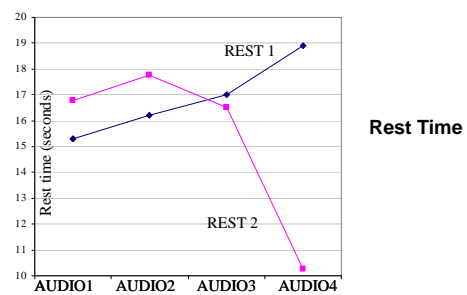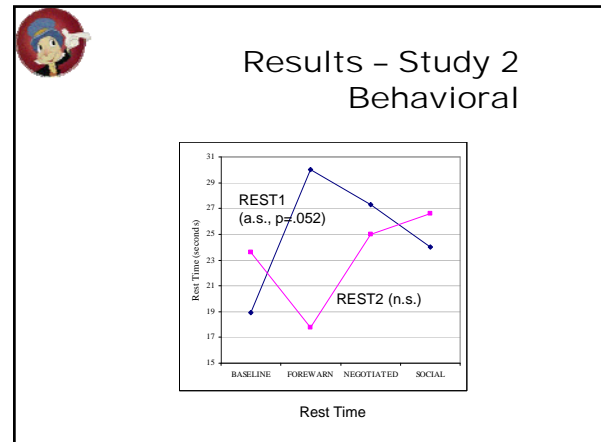
Hypotheses

Short-term Compliance vs Politeness
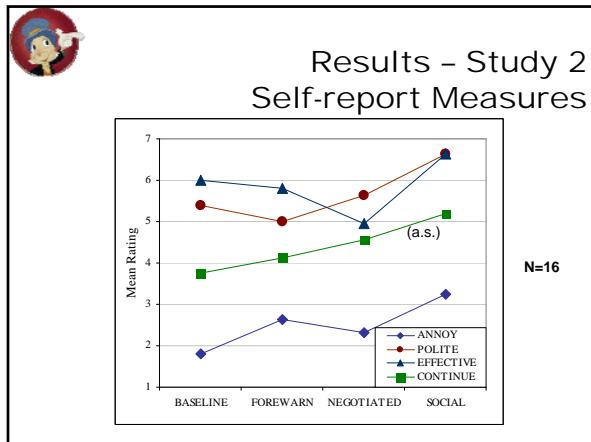
Long-term Usage vs Politeness

Long-term Adherence vs Politeness

Application: Wrist Rests

Results – Study 1
Self-report Measures

POLITE

CONTINUE

ANNOYING

N=29

AUDIO1 | AUDIO2 | AUDIO3 | AUDIO4

Results – Study 1
Behavioral

REST 1

REST 2

Rest time (seconds)

Rest Time

AUDIO1 | AUDIO2 | AUDIO3 | AUDIO4

Results – Study 2
Self-report Measures

N=16

Results – Study 2
Behavioral

Rest Time

Wearable Agent Field Study in Process

- Primary hypothesis: real-time intervention more effective than retrospective.
- 5-week, 5-treatment within-subjects design
- 100 free-living, sedentary adults

Conclusion

- Relational bonding is important in long-term interactions

Plug: CHI Engagement by Design workshop
ebd.wikispaces.com

bickmore@ccs.neu.edu
relationalagents.com