Judgmental Bot:
Conversational Agents for Online Mental Health Screening

Alan R. Dennis

With M. Rahimi, A. Kim, and S. Ayabakin

Outline

Background
Study 1
Study 2
Study 3
Implications

Background

Mental Health

Mental health epidemic

50% of the US population
will face mental health issues
   5% need long term care
    15% need short term care
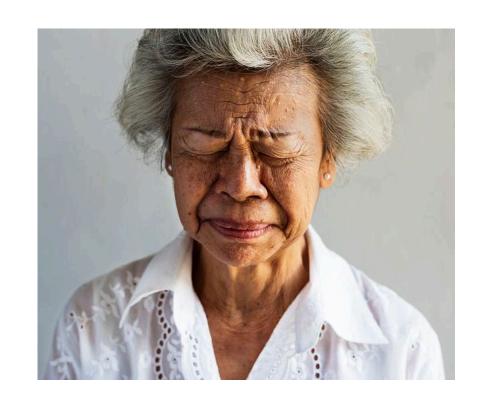
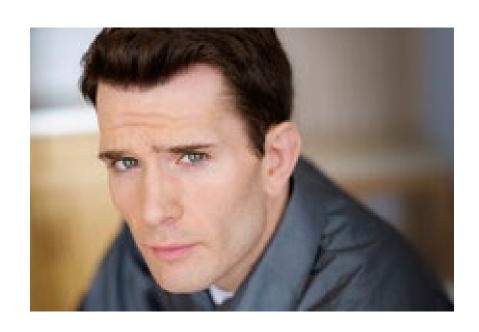Shortage of mental health
professionals

Stigma

Negative attitudes or discrimination towards those with mental health issues

Fear of Judgmentalness

is a barrier
to seeking care
to disclosing information

Mental Health Screening

PHQ 9

Over the last *2 weeks*, how often have you been bothered by any of the following problems?
(use "✓" to indicate your answer)

	Not at all	Several days	More than half the days	Nearly every day
1. Little interest or pleasure in doing things	0	1	2	3
2. Feeling down, depressed, or hopeless	0	1	2	3
3. Trouble falling or staying asleep, or sleeping too much	0	1	2	3
4. Feeling tired or having little energy	0	1	2	3
5. Poor appetite or overeating	0	1	2	3
6. Feeling bad about yourself—or that you are a failure or have let yourself or your family down	0	1	2	3
7. Trouble concentrating on things, such as reading the newspaper or watching television	0	1	2	3
8. Moving or speaking so slowly that other people could have noticed. Or the opposite — being so figety or restless that you have been moving around a lot more than usual	0	1	2	3
9. Thoughts that you would be better off dead, or of hurting yourself	0	1	2	3

add columns   +   +

(Healthcare professional: For interpretation of TOTAL, please refer to accompanying scoring card).   **TOTAL:**

AI & Chatbots

Good for structured data
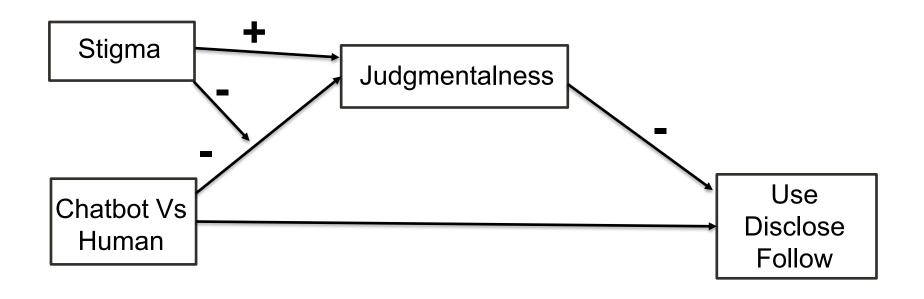
Always available

No moral judgment

No prior research on judgmentalness

Most research has assumed less judgmental

Research Model

Study 1

Participants

Who contacts mental health screening services?

80% need no treatment
15% need short term care
5% need long term care

US Population

80% need no treatment
15% need short term care
5% need long term care

300 participants on Amazon Mechanical Turk

49% female, mean age 43
55% prior experience with mental health issues
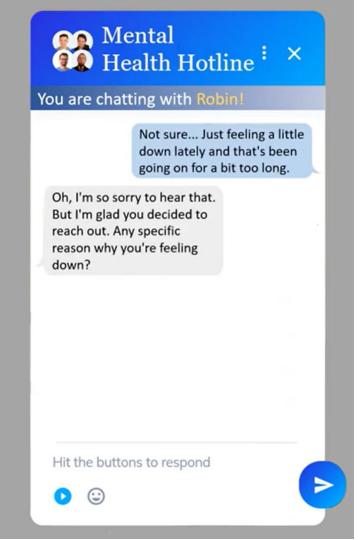85% were White, 8% Black, 6% Asian, and 1% other

KELLEY SCHOOL OF BUSINESS
INDIANA UNIVERSITY

Vignette Experiment

Participants watched a video of a screening following best practice

Half told it was a nurse
Half told it was AI

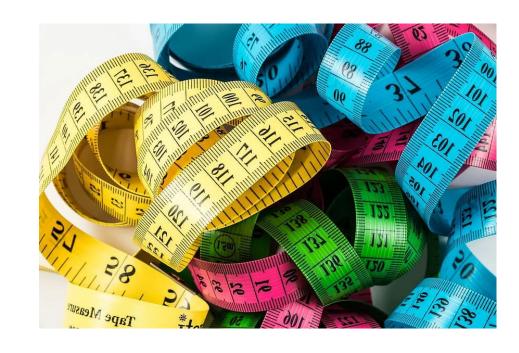Complete control to ensure all treatments were exactly the same

Measures

Stigma
Willingness to Use
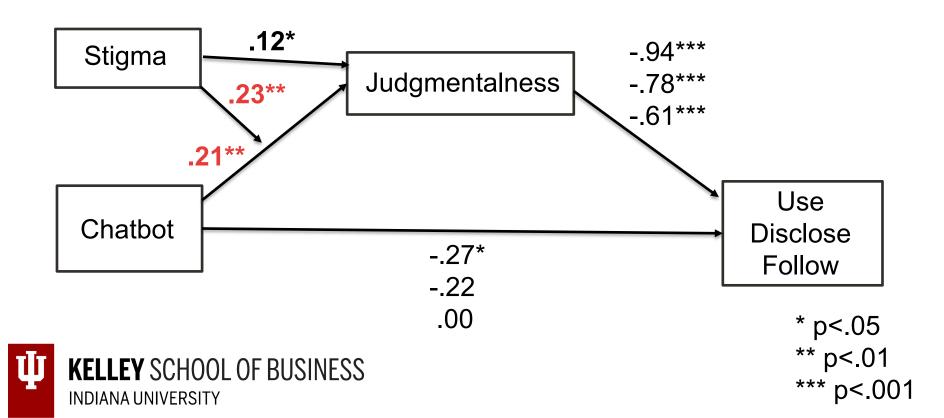Disclose Information
Follow Advice

Judgmentalness
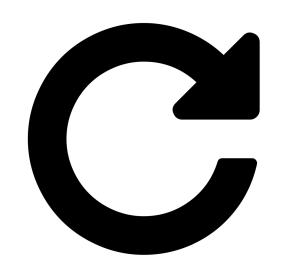
Controls
age, gender, race, income, past mental health experience, trust in hospital

Results

Stigma →[**.12***] **Judgmentalness**

Stigma →[**.23****] Use Disclose Follow

Chatbot →[**.21****] Judgmentalness

Judgmentalness →[-.94*** / -.78*** / -.61***] Use Disclose Follow

Chatbot →[-.27* / -.22 / .00] Use Disclose Follow

* p<.05
** p<.01
*** p<.001

KELLEY SCHOOL OF BUSINESS
INDIANA UNIVERSITY

Robustness Check

New Treatment with
No emotional support

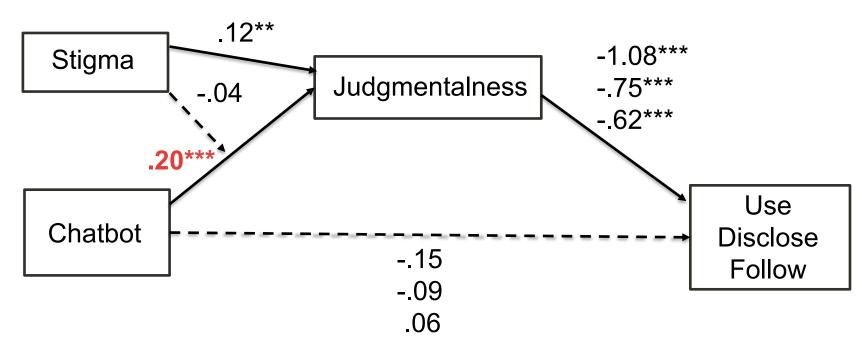New Measures
of Stigma and
Judgmentalness

541 participants from
Prolific Academic

Results

New Treatment
(**No** Emotional Support)

Stigma → Judgmentalness: .12**

Stigma ⇢ Use Disclose Follow: -.04

Chatbot → Judgmentalness: **.20***

Judgmentalness → Use Disclose Follow:
-1.08***
-.75***
-.62***

Chatbot ⇢ Use Disclose Follow:
-.15
-.09
.06

KELLEY SCHOOL OF BUSINESS
INDIANA UNIVERSITY

Study 2

Qualitative Interview Study

Goal was to understand why people perceive chatbots to be judgmental

Specifically, the theoretical factors influencing **chatbot judgmentalness**

Participants

41 Participants from Indiana University
Students (34), Faculty (6), and Staff (1)

59% female, mean age 26
93% prior experience
35% were White, 8% Black, 40% Asian, and 15% other

Method

Watched the Original Video
Answered Interview Questions about
Why the chatbot might be perceived as judgmental

Analysis

Two researchers, one interviewer, one non-interviewer

Open Coding

Thematic Coding

Theoretical Coding

Diagramming Relationships

KELLEY SCHOOL OF BUSINESS

INDIANA UNIVERSITY

Table 4. Factors Influencing Perceptions of Judgmentalness in Study 2

First Order Codes	Second Order Codes
Chatbot incapable of experiencing human real-life experience of the patient Chatbot incapable of connecting its own personal experiences to patient's problems Another human with experiences to understand the patient's problems is comforting	Real-Life Experience
Chatbot lacking a deep understanding of the patient's problem Chatbot and the patient lacking the same deep understanding of the patient's problems Chatbot lacks a deep understanding of how patient's problems have affected them	Understanding: *Deep Understanding*
Chatbot summarizing the patient's problem means understanding Chatbot asking follow-up questions means understanding Chatbot sharing related stats shows understanding	Understanding: *Surface Understanding*
Chatbot not being able to understand human emotions Chatbot cannot feel emotions Chatbot lacks qualia (internal/subjective experiences arising from sensory stimulation)	Understanding: *Emotional Understanding*
Chatbot is incapable of understanding human social experiences Chatbot cannot understand social consequences of patient's problems Chatbot cannot understand how other people might react to the patient's problems	Understanding: *Social Meaning Understanding*
Chatbot incapable of conveying a sense of being heard Chatbot incapable of providing human connection and companionship Patients need to feel someone is there for them	Validation: *Being Heard*
Chatbot not being able to make the patient feel like their problem are valid Chatbot acknowledging that the feelings of the patient are important Perception of being heard is necessary to validating the patient's problems	Validation: *Being Validated*

Table 5. Study 2 Example Quotations

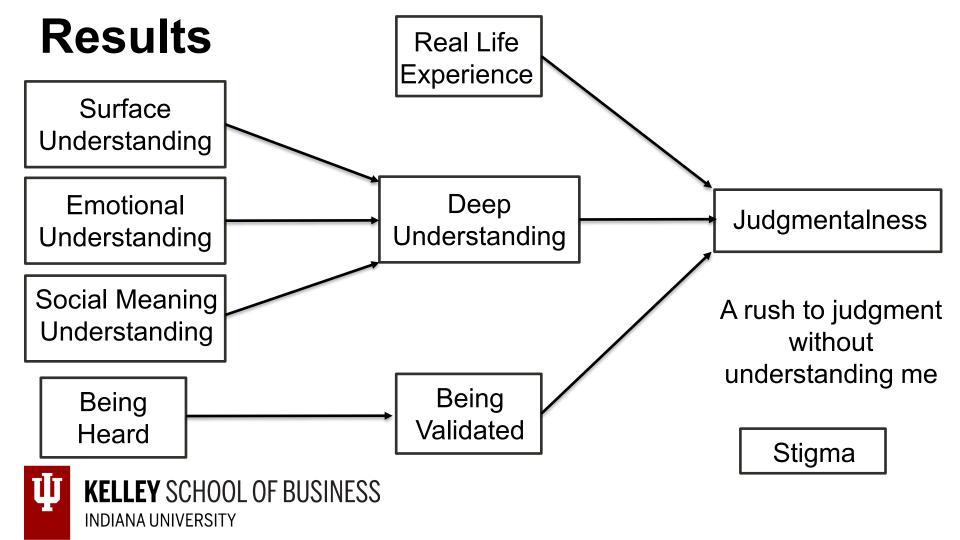| Codes	Example Quotations
Real-Life Experience	"For example, if I'm talking to someone, then he can tell exactly I've seen this kind of situation with one of my patients in the last two weeks and he was having [a] similar experience… [humans] are much better compared to the things that robots can show."
Deep Understanding	1. "It just feels like, yeah, you're talking to an algorithm. You want someone else to understand your pain and to feel that pain." 2. "Kind of talking to [the] chatbot… just picking up text strings or something and… didn't feel like it knew anything…[but not] kind of like genuine recognition that, hey, this person is saying that they're feeling bad…"
Surface Understanding	1. "…I felt like it was a little bit surface level just because yeah, they were listing the issues and the AI was kind of giving very general replies to them…" 2. "Robots can certainly understand the thing, because it's following the usual semantics that you[r] robots are usually trained on."
Emotional Understanding	1. "You're talking to something that's just trying to pick up certain phrases. It doesn't feel [like] the actual empathy you want to." 2. "…here we can see that humans can certainly show better empathy compared to [chatbot]".
Social Meaning Understanding	"…the chatbots don't really have like the nuance that we do in terms of these types of situations just because it's not like a real human and doesn't exist [in the] society that we do."
Being Heard	"I just think we just seek human connections and we just want to feel like there's someone out there listening to us. And I think that if we think that it's a robot…it doesn't feel like that as much."
Being Validated	1. "…[the chatbot] making me feel like my problem is minimized or not being validated, acknowledged that could be an example of being judgmental." 2. "… [if] the person was not feeling well and acknowledgement of the feeling, not just the fact that [it] was disclosed. I do think that if there's not an acknowledgement of the feeling that would be detrimental."

Results

Real Life Experience

Surface Understanding

Emotional Understanding

Social Meaning Understanding

Deep Understanding

Judgmentalness

A rush to judgment without understanding me

Being Heard

Being Validated

Stigma
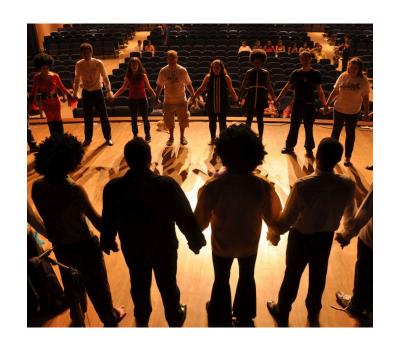
Study 3

KELLEY SCHOOL OF BUSINESS
INDIANA UNIVERSITY

Two Experiments

Study 3a Cloud Research

413 Participants

56% female, mean age 47

62% prior experience

80% were White, 11% Black, 4% Asian, and 5% other

Study 3b IU Students

393 Participants

58% female, mean age 19

43% prior experience

63% were White, 1% Black, 31% Asian, and 5% other

KELLEY SCHOOL OF BUSINESS

INDIANA UNIVERSITY

Methods

Treatments same as Study 1

Judgmentalness and three outcomes
Same as Study 1 Robustness Check

Developed and Validated
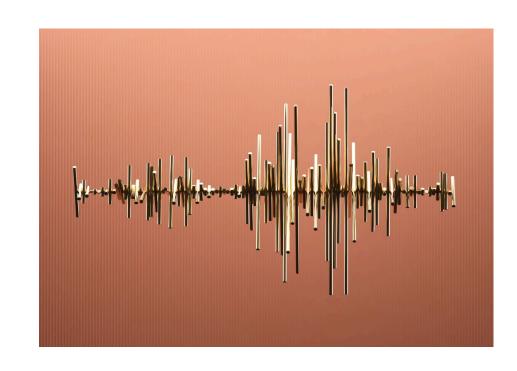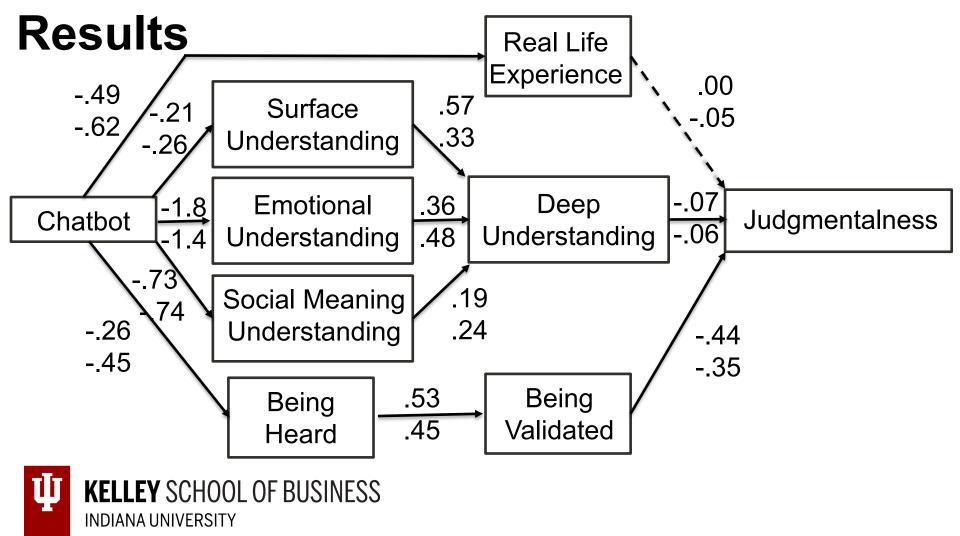new Measures for the
other constructs

Results

Chatbot more judgmental

Study 3a p=.022, d=.19
Study 3b p=.031, d=.16

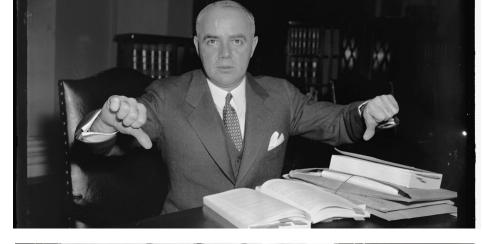KELLEY SCHOOL OF BUSINESS
INDIANA UNIVERSITY

Results

Implications

Judgmentalness
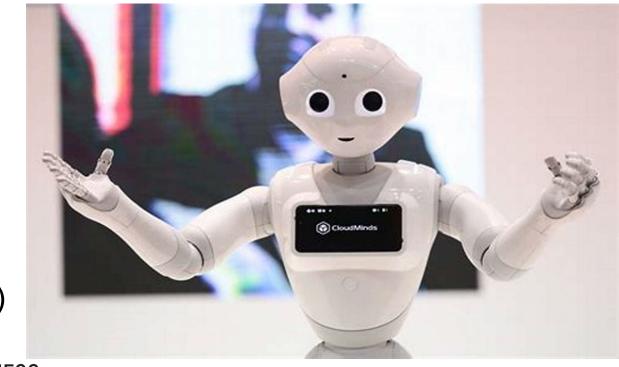
Moral judgment

A rush to judgment without understanding

AI is More Judgmental

It doesn't have a deep understanding (Lesser Importance)

People don't feel heard and validated (Greater Importance)

Human and AI Behavior was Identical

Making the AI better will have only a small effect

Emotional Support

Research Implications

Do not theorize that AI is less judgmental

More research on judgmentalness

More research on AI and emotional support

Practical Implications

Should we use AI for
Mental Health Screening?

Chatbot with
Emotional Support

Online Form

Questions