
# Human Factors Methods for Digital Design

Allison Schmidt, MS

Yeganeh Shahsavar, PhD

MAY 1-5, 2026 · DSAI MT. WASHINGTON, BALTIMORE

# Introduction

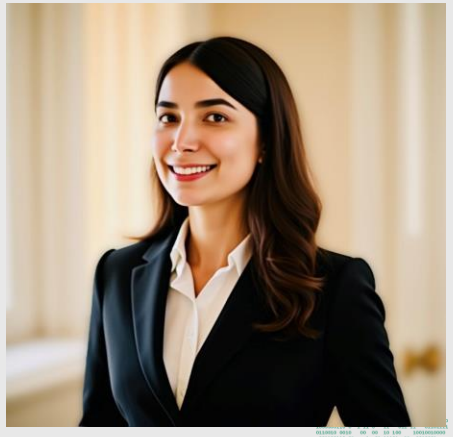

## Allison Schmidt, MS

- Human Factors Engineer
- Armstrong Institute for Patient Safety and Quality
- Johns Hopkins Medicine



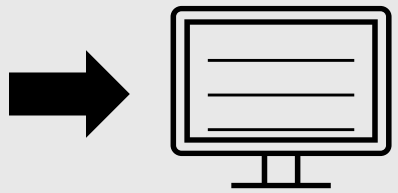
## Yeganeh Shahsavar, PhD

- Postdoctoral Research Fellow
- Armstrong Institute for Healthcare Human Factors
- Johns Hopkins Medicine

# Background: Patients message ambulatory clinics and providers

Patient send messages to care team



Hundreds of messages are received in a clinic 'pool'

**Messages are addressed by different roles in the clinic with different expertise**

**Medical Office Coordinator**

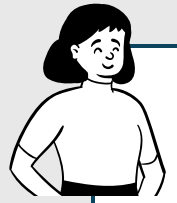
=

Non-medical questions

**Nurse**

=

Routine medical questions

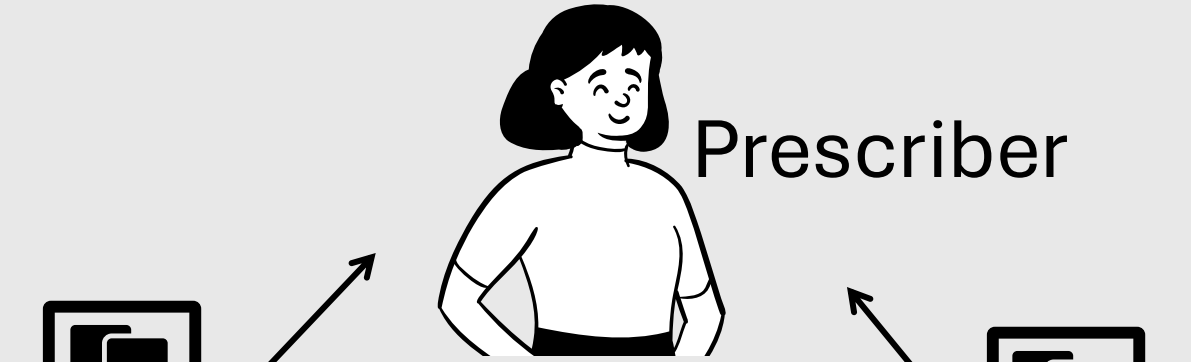


**Prescriber**

=

Complex medical questions

# Problem: Messages are reviewed by multiple people before they reach the person that can address it



Prescriber

Burden reading messages they cannot address

Delays from re-routing messages

Medical Office Coordinator

Nurse

Flourish-A-Thon

# Solution: Artificial intelligence (AI) automates patient message triage

Natural Language Processing analyzes incoming message text

→

Category is assigned to message

- Clinical
- Clinical (probable)
- Administrative (probable)
- Administrative

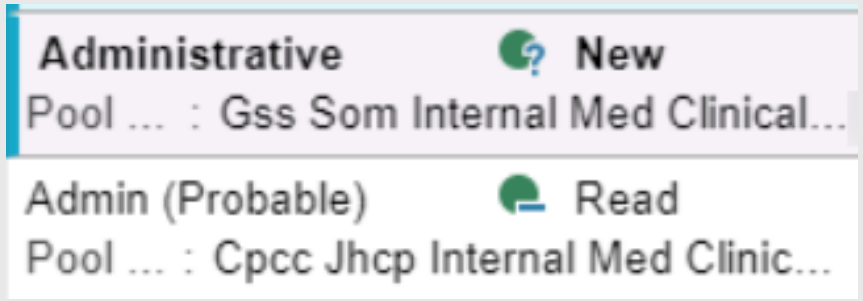
→

Clinic roles assigned to review messages in a specific category

## Example

**Subject**: Check-up Can I schedule an appointment with Dr A?

Natural Language Processing of message




**Category Assigned:** Administrative

Medical Office Coordinator addresses message with "Administrative" category

Flourish-A-Thon

# Objectives

✓ Objective #1: Does the patient message categorization tool accurately categorize incoming patient messages?

? Objective #2: Is the patient message categorization tool…

1) Usable – Easy to learn, understand, navigate, and use
2) Useful – Meet user needs and improves performance and productivity
3) Trustworthy – Perceived to be reliably accurate

# Usability Analysis

1. Understand users and their current work
2. Perform expert evaluation using design heuristics
3. Aggregate design flaws and assign severity score
4. Continuously improve design

| User Control and Freedom |
| --- |
| Users cannot recategorize messages, causing workarounds to get messages to the right responder |

| Visibility of System Status |
| --- |
| The original message category persists. As conversation evolves, the category is less useful and noisy |

# What we Learned

Technically functional and accurate AI Tool

≠

Used and trusted AI tool that *fits user needs*

***Incorporate Human Factors Engineering Early***
to proactively design AI tools that fit the user needs and yield successful AI tool designs

Flourish-A-Thon

# What is Human Factors Engineering?

Flourish-A-Thon

# Fit the Design to the Person, Not the Person to the Design

**Understand:**
Interactions among people and other elements of a system

**Develop and Apply:**
Theory, principles, data and methods of ergonomics

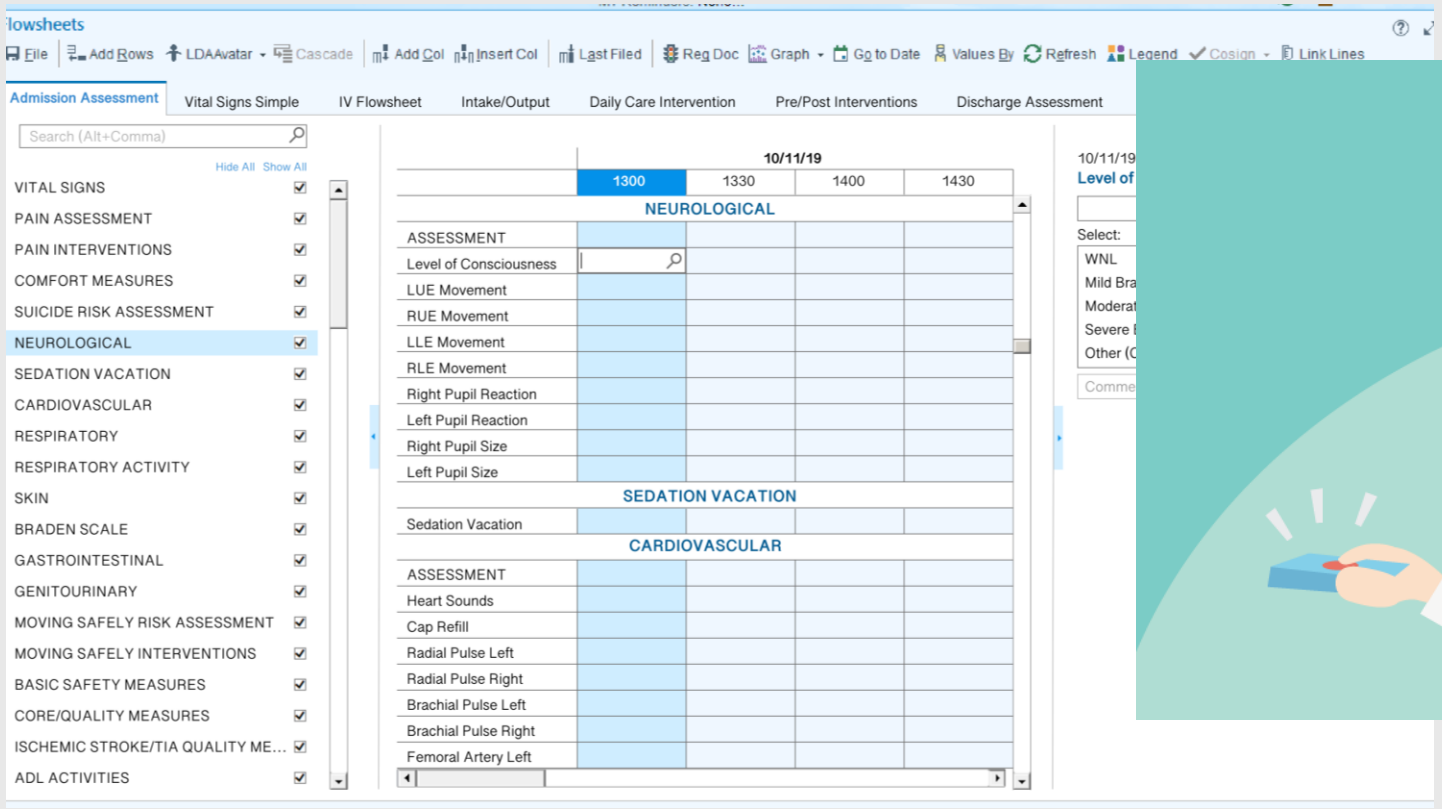
**Design and Evaluate:**
Tools, systems, tasks, jobs, and environment

**Improve:**
Human Wellbeing (less errors, fatigue, stress and injuries)
System Performance (faster, more effective, and more efficient)

# What does 'Fit' mean?

## Micro-Level

Phone, health record, device




## Macro-Level

Device used in the context like home, hospital, school, etc.

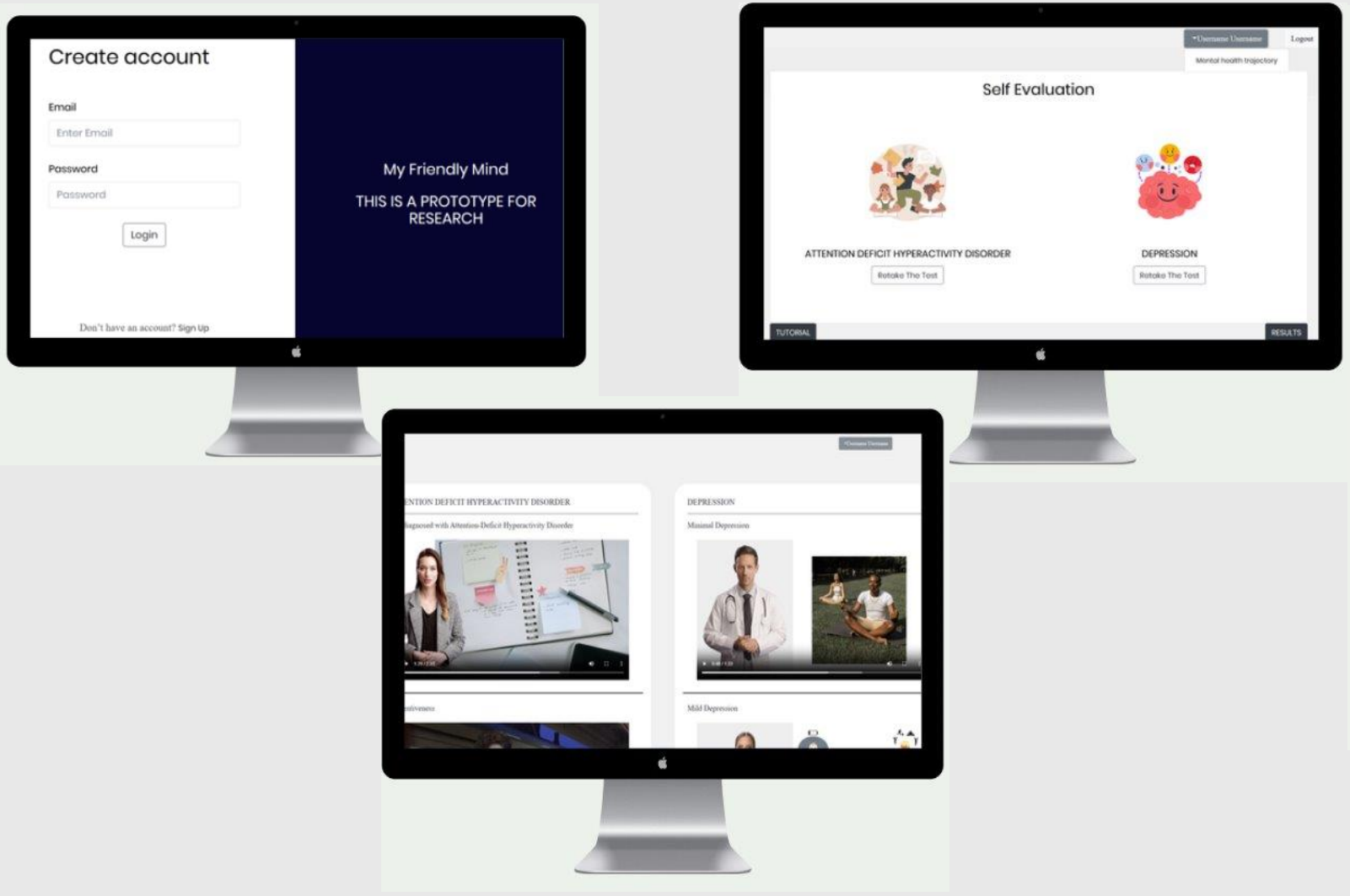
Flourish-A-Thon

# Case Study:

## Application of Human Factors Methods for a Mental Health App Evaluation

Flourish-A-Thon

# My Friendly Mind App




**Interface:**

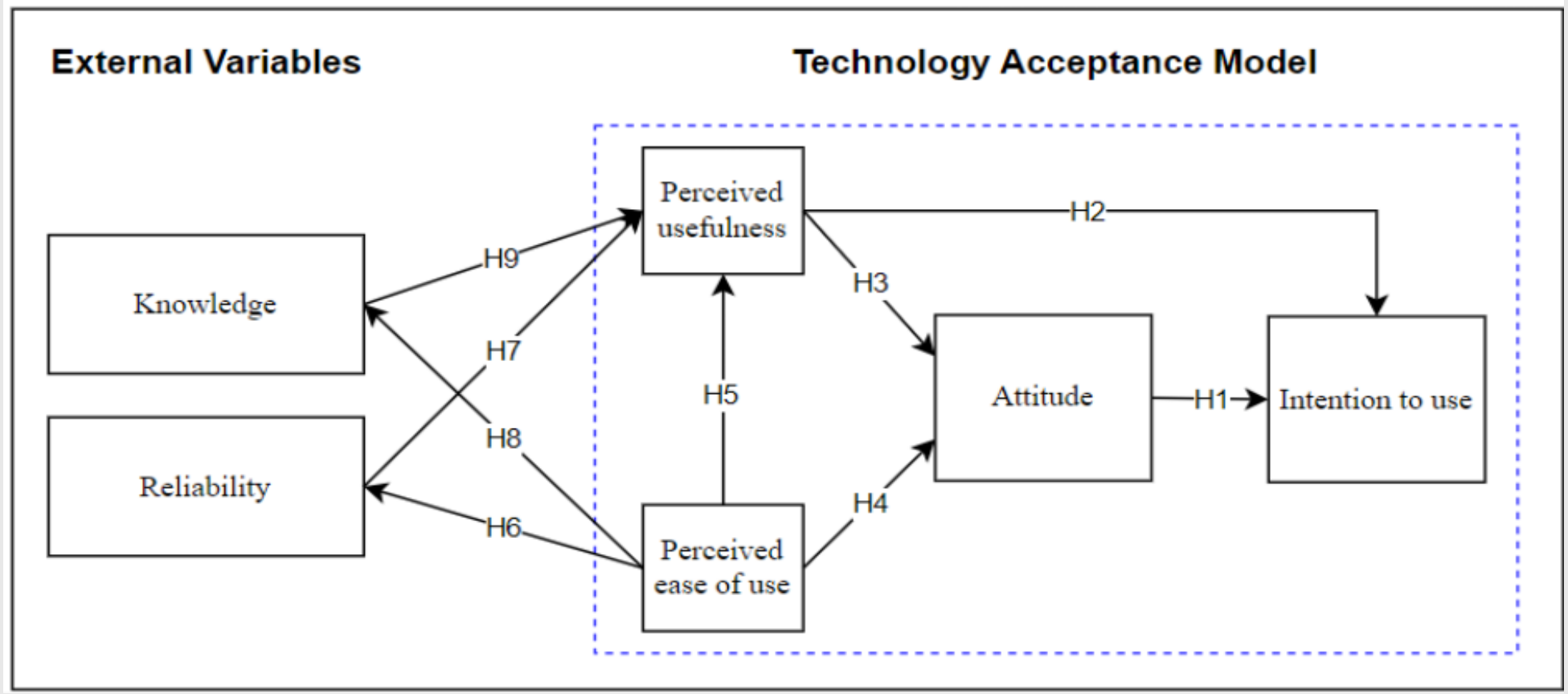
- Sign up page
- Login page
- Home Page
- **Self-report process**
- Outcome page
- Result page (Videos)
- Tutorial page (Videos)

Based on two clinically validated tools:

- Patient Health Questionnaire-9 (PHQ-9)
- Adult ADHD Self-Report Scale (ASRS)

Flourish-A-Thon

# Study Hypotheses Informed by the Technology Acceptance Model (TAM)




Article: Assessing user acceptance of a mental health app & its impact on depression and attention deficit hyperactivity disorder related knowledge A mixed method experimental study.pdf

# Data Collection

Phase 1: → Phase 2: → Phase 3:

### Baseline Survey

- Pre-Knowledge Survey: Our custom-designed questionnaire assessing knowledge of depression and ADHD.
- Includes 3 multiple choice 1 yes, no and 2 open-ended questions.

### Post Intervention Surveys

- Mobile App Rating Scale (MARS)- **23 questions**
- System Usability Scale (SUS)- **10 questions**
- Health Information Technology Usability Evaluation Scale (Health-ITUES)- **20 questions**
- Post-knowledge survey
- Trust Survey- **5 questions**

  5 point Likert's scales 'not at all' to 'a lot'.

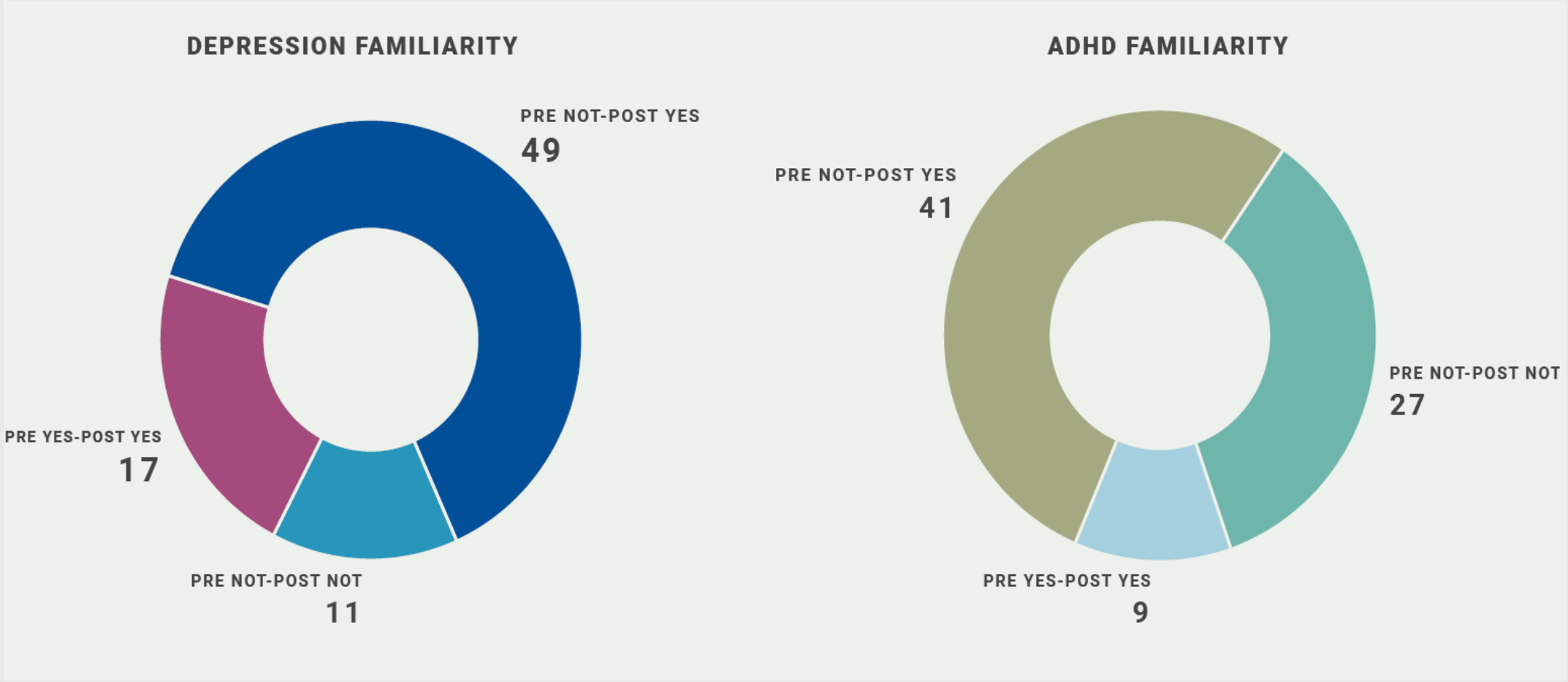
### Post Intervention Interview

1. Experience using the app.
2. App's potential to help with depression or ADHD.
3. Thought process in trusting AI or doctor outcomes.
4. Likes and dislikes (list 3-5 features each).
5. Insights gained from videos.
6. Suggestions for app improvement

### Survey's links:

Knowledge:
Knowledge.docx
MARS:
MARS.docx
SUS:
SUS.docx
Health-ITUES:
Health-ITUES.docx
Trust:
Trust.docx

# Baseline and Post-Intervention Knowledge Survey Findings




Flourish-A-Thon

# Post Intervention Survey Findings

| | Category | Mean Score | Benchmark |
|---|---|---|---|
| **System Usability Scale (SUS)** | Usability | 70.62 | >70 Moderate [4] |
| **Mobile App Rating Scale (MARS)** | Engagement | 3.48 | Acceptable [5] |
| | Functionality | 4.60 | Excellent [5] |
| | Aesthetic | 4 | Good [5] |
| | Information | 3.37 | Acceptable [5] |
| | App Quality | 3.86 | Acceptable [5] |
| | App Subjective Quality | 3 | Acceptable [5] |
| **Health Information Technology Usability Evaluation Scale (Health-ITUES)** | Usability & Effectiveness | 3.88 | Moderate [6] |

# Post Intervention Interview Findings

### Experience with the App:

- User-friendly, straightforward, and helpful as a self-assessment tool.
- Videos were valued for providing mental health management tips.

**P2** "THE APP WAS GOOD. IT WAS USER FRIENDLY. I DIDN'T HAVE ANY PROBLEM WITH IT."

**P51** IT WAS GOOD, HELPFUL, AND VERY INFORMATIVE. I ENJOYED IT. I LEARNED MORE.

### Trust in Outcome:

- Personal experiences
- Perceived symptom relevance.

### Liked Features:

- Straightforward & user-freindly

### Disliked Features:

- Unnatural AI avatars in videos.

### Suggestions for Improvement:

- Complexity and Depth of Questions

**P10** "DEPRESSION IS SUCH A COMPLEX TOPIC... YOU WOULD NEED TO ASK MORE QUESTIONS.

**P39** "ASK A COUPLE MORE QUESTIONS BEFORE THEY DIAGNOSE IT."

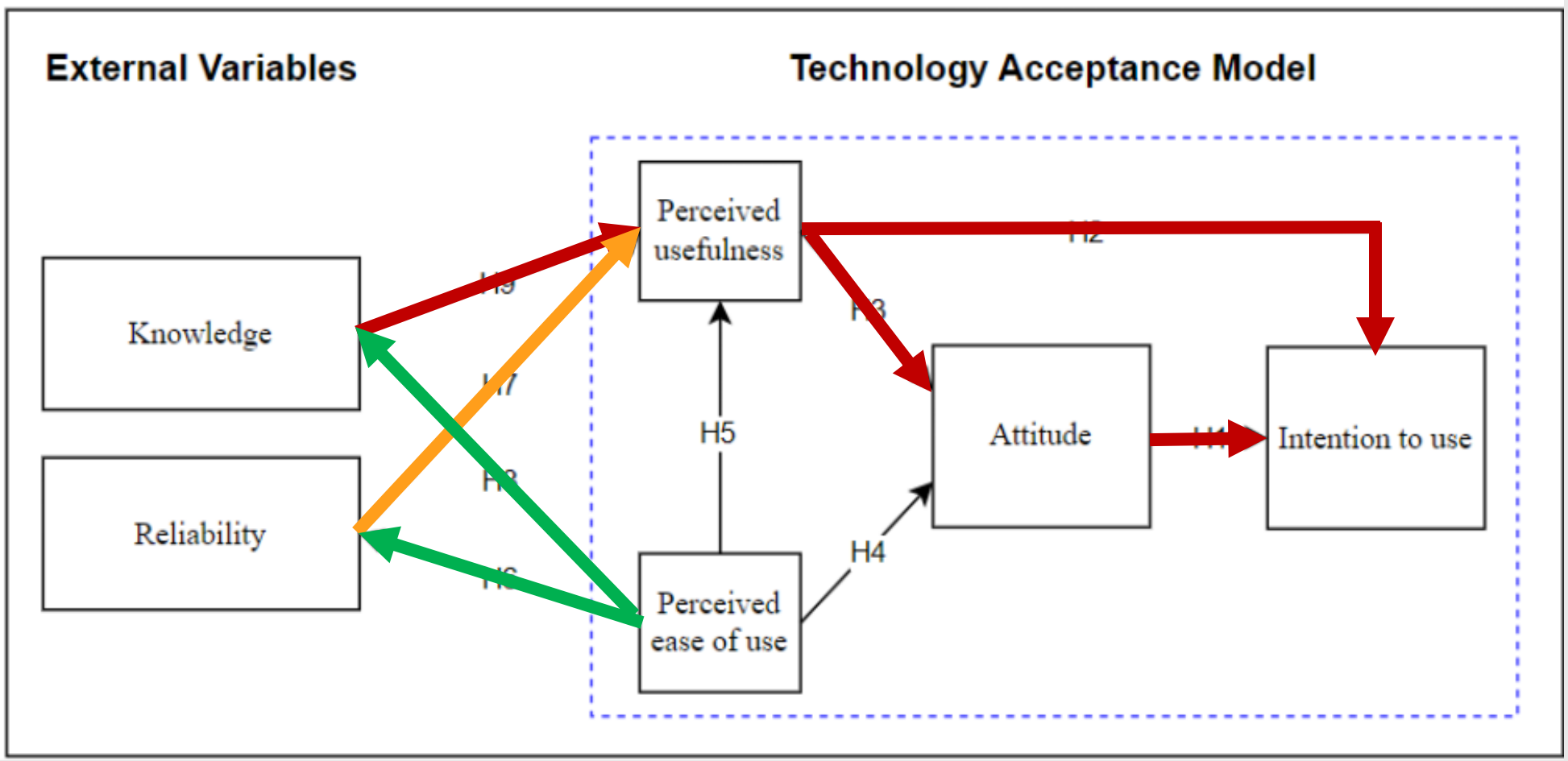
- Video Speakers

**P30** "YOU DON'T WANT TO LISTEN TO A ROBOT."

**P26** "THE LADY SPEAKING WAS A LITTLE CREEPY IN THE VIDEOS."

Flourish-A-Thon

# Technology Acceptance Model (TAM) Finding




Framework Questions.png

Flourish-A-Thon

# Measuring Usability Using Physiological Factors

- Eye movements
- Pupil size
- Blink rate
- Heart rate
- Galvanic skin response
- Brain signals

# Human Factors Evaluation Questions to Consider

Flourish-A-Thon

# Digital Tool Evaluation

## Useful

- Who is the end user?
  - What are each user's needs?
- What problem does this tool address?
- How does it improve the following:
  - Productivity
  - Efficiency
  - Performance
  - Accuracy
  - Wellbeing

## Usable

- Is the tool easy to learn?
- Is the tool easy to use?
- Is the tool easy to understand?
- Is the tool use flexible?
  - Can users 'go back'
  - Do users have freedom to interact how they wish
- Does the tool prevent errors? Recover from errors?

## Trustworthy

- Is the output reliably accurate?
- Does the tool explain the rational for the output?
- Is the output understandable?
- Is the information private and secure? How do users know?