

**Medship: Affective Computing for Building
Empathetic Behaviors Toward Patients with
Substance Use Disorders**

by

Caleb M. Harris

S.B. Computation and Cognition, Massachusetts Institute of
Technology (2022)

Submitted to the Department of Electrical Engineering and Computer
Science and the Department of Brain and Cognitive Sciences
in partial fulfillment of the requirements for the degree of

Master of Computation and Cognition

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

September 2022

© Massachusetts Institute of Technology 2022. All rights reserved.

Author
Department of Electrical Engineering and Computer Science and the
Department of Brain and Cognitive Sciences
August 05, 2022

Certified by
Rosalind W. Picard
Professor of Media Arts and Sciences
Thesis Supervisor

Accepted by
Mehrdad Jazayeri
Director of Education, the Department of Brain and Cognitive Sciences

Medship: Affective Computing for Building Empathetic Behaviors Toward Patients with Substance Use Disorders

by

Caleb M. Harris

Submitted to the Department of Electrical Engineering and Computer Science and the Department of Brain and Cognitive Sciences on August 05, 2022, in partial fulfillment of the requirements for the degree of Master of Computation and Cognition

Abstract

Opioid use is on the rise, with overdose deaths quadrupling since 1999 [34]. Consequently, so is substance use disorder (SUD), an illness caused by repeated use of substances such as alcohol or drugs that result in clinically significant impairment. Although physician interactions with patients with SUDs are dramatically increasing in frequency, the majority of medical training still fails to address the importance of building empathy and minimizing stigma in such clinical interactions. Furthermore, physicians receive only minimal instruction regarding the expression of empathy and its role in building rapport and eliciting positive responses from patients with SUDs. Such strategies not only improve the immediate clinical interaction by contributing to a warm, stigma-free environment, but also improve the long-term outcomes of the patient by driving them toward care instead of away from it. This thesis both identifies the affective features and expressions most attributed to positive clinical perception, from the perspective of actual patients with SUDs, and introduces Medship, a web-application-based tool embedded with affective computing models to provide real-time affective training for medical school students and physicians in simulated clinical interactions with patients with SUDs.

Thesis Supervisor: Rosalind W. Picard

Title: Professor of Media Arts and Sciences

Acknowledgments

First and foremost, I would like to acknowledge my thesis supervisor, Professor Rosalind Picard, for providing me the opportunity to do this research in the Affective Computing Group of the Media Lab. Roz has been a tremendous help during the past year— from guiding me to this collaborative project based on my interests, to aiding me in the design and development and to assisting me along the way as I encountered roadblocks. I am very grateful to Roz, for without her, none of this would have become a reality.

Furthermore, I would like to acknowledge Matt Groh, my UROP advisor who gave me the initial opportunity to work in the Affective Computing Group. Matt has been an incredible advisor not only during my UROP as an undergraduate, but also throughout my time as a graduate student. During our time working together, we co-authored two papers, which is something I would have never expected from undergraduate research. Through these experiences Matt challenged me in ways that technically allowed me to expand my skill set beyond what I could learn from coursework.

I would also like to acknowledge my collaborators from Weill Cornell Medical College, Dr. Jonathan Avery and Giselle Appel, who brought the initial idea to us and were instrumental in securing funding and carrying out the design and execution of the clinical trial. Furthermore, Ila Kumar, another graduate student in the Affective Computing Group, has helped tremendously throughout the year.

Additionally, I would like to acknowledge my two academic advisors during my four years at MIT, Dr. Joshua Tenenbaum and the late Dr. Patrick Winston. During the short time I knew him, Dr. Winston was instrumental in guiding me through a difficult transition into MIT's rigor. He helped me realize my passion for computer science, and importantly how technology can be used to improve lives. Dr. Tenenbaum similarly helped me shape my interests in the cross section of computation and cognition, a field in which I am now deeply motivated.

I would also like to acknowledge my partner, Micah, who has been with me

throughout the entirety of my time at MIT, and been there for me during the roller coaster of ups and downs. I can not even begin to imagine how challenging it would have been to make it to this point without her, and I am forever grateful.

Finally, I would not have made it through MIT without the support of my friends and family. Particularly, my fraternal brothers of Phi Kappa Theta have given me a strong sense of community and a place I now call home. The community has driven me to succeed, supported me when I failed and been there for me in times of need, and I am confident that many of the relationships I have formed will be life long. Lastly, thank you to my family members for the continual love and care. Each of you have helped me become the person I am today, and supported me always.

Contents

1	Introduction	13
1.1	Motivation	13
1.2	Background	14
1.2.1	Addiction and Stigma	14
1.2.2	Physician Attitudes and Education	15
1.2.3	Marginalized Populations	16
1.3	Thesis Overview	16
1.3.1	Part 1: Interviews	17
1.3.2	Part 2: Raters	18
1.3.3	Medship	18
1.3.4	Impact	19
2	Related Work	21
2.1	Virtual Training	21
2.2	Coachware	22
3	Study Design	23
3.1	Description	23
3.2	Part 1: Interviews	24
3.2.1	Setup	24
3.2.2	Task 1	25
3.2.3	Task 2	25
3.2.4	Task 3	26

3.2.5	Task 4	27
3.2.6	Task 5	27
3.3	Part 2: Raters	27
4	System Design	31
4.1	Overview	31
4.2	Expression Practice	33
4.3	Training	34
5	Results & Discussion	39
5.1	Clinical Study	39
5.2	Medship	42
5.3	Discussion	42
6	Future Work	45
6.1	Continuation of Medship	45
6.2	AI-Based Coachware	46
A	Figures	47
B	Tables	65

List of Figures

A-1	Group 1 - Physicians: Demographics - Age and identity	48
A-2	Group 1 - Physicians: Demographics - LGBTQ+ and race	49
A-3	Group 1 - Physicians: Demographics - Years practicing and medical specialty	50
A-4	Group 1 - Physicians: Personal experience with addiction and aspects deemed most important to see in a physician with respect to communicating emotional expression	51
A-5	Group 2 - Patients with SUD: Demographics - Age and identity . . .	52
A-6	Group 2 - Patients with SUD: Demographics - LGBTQ+ and race . .	53
A-7	Group 2 - Patients: Aspects deemed most important to see in a physician with respect to communicating emotional expression	54
A-8	Medship: Landing	55
A-9	Medship: Home	56
A-10	Medship: About	57
A-11	Medship: Expression practice landing	58
A-12	Medship: Expression practice example	59
A-13	Medship: Training landing	60
A-14	Medship: Training setup	61
A-15	Medship: Training example	62
A-16	Medship: Background information	63

List of Tables

5.1	Group 3 - Raters with SUD: Summary of the raters' perceptions of physicians' responses in clips from part 1	40
5.2	Group 3 - Raters with SUD: Metrics of inter-rater reliability for raters' perceptions of physicians' responses in clips from part 1	40
5.3	Group 3 - Raters with SUD: Agreement between raters on physicians' responses in clips from part 1	41
5.4	Group 3 - Raters with SUD: Percentage agreement between raters on physicians' responses in clips from part 1	41
5.5	Group 3 - Raters with SUD: Summary of the raters' perception of the physicians, taking the perspective of the patients	42
B.1	Group 3 - Raters with SUD: Demographics - Age	65
B.2	Group 3 - Raters with SUD: Demographics - Identity	66
B.3	Group 3 - Raters with SUD: Demographics - LGBTQ+	67
B.4	Group 3 - Raters with SUD: Demographics - Race	68
B.5	Group 3 - Raters with SUD: Demographics - Opioid use status	69
B.6	Group 3 - Raters with SUD: Demographics - Years with an active SUD	70
B.7	Group 3 - Raters with SUD: Demographics - Years in recovery from SUD	71
B.8	Group 3 - Raters with SUD: Survey 1 - Agreement between raters on physicians' responses in clips from part 1	72
B.9	Group 3 - Raters with SUD: Survey 2 - Agreement between raters on physicians' responses in clips from part 1	73

B.10 Group 3 - Raters with SUD: Survey 3 - Agreement between raters on physicians' responses in clips from part 1	74
B.11 Group 3 - Raters with SUD: Survey 4 - Agreement between raters on physicians' responses in clips from part 1	75
B.12 Group 3 - Raters with SUD: Survey 1 - Raters' perception of the physicians, taking the perspective of the patients	76
B.13 Group 3 - Raters with SUD: Survey 2 - Raters' perception of the physicians, taking the perspective of the patients	77
B.14 Group 3 - Raters with SUD: Survey 3 - Raters' perception of the physicians, taking the perspective of the patients	78
B.15 Group 3 - Raters with SUD: Survey 4 - Raters' perception of the physicians, taking the perspective of the patients	79
B.16 Group 3 - Raters with SUD: Survey 5 - Raters' perception of the physicians, taking the perspective of the patients	80

Chapter 1

Introduction

1.1 Motivation

Substance overdose deaths in the U.S. topped 100,000 annually, for the first time ever, for the 12 month period ending in April 2021. It is estimated that of those deaths, 75,673 were overdose deaths from opioids, a 35% increase of that from the year prior [33]. The modern-day patient-physician relationship in substance use disorder (SUD) is flawed due to widespread stigma toward patients, yet the healthcare field has not created engaging and targeted tools to help future physicians decrease their biases. Physician attitudes toward patients with a SUD are more negative than attitudes toward patients with other medical and psychiatric conditions, and these biases worsen over the course of one's career [13][4][14][15][32]. Furthermore, individuals with a SUD from marginalized populations face compounded discrimination [16][41][36][39]. This stems from both a lack of physician knowledge about SUD and patients' tendency to relapse care [46]. As a result, physicians' communication, display of empathy, and involvement with patients deteriorates [9][30][5]. Nevertheless, the majority of medical training fails to improve physician awareness and management of affective expressions toward patients with a SUD [38][25][23].

This thesis introduces Medship, a platform and technology aimed at helping medical trainees both understand their own outward expressions and practice contextually-relevant clinical situations in which affective expression, engagement and empathy

building are paramount. Medship is designed with the goal of improving empathetic communication and behaviors, and relies on the following principles:

1. Empathy is malleable.
2. Empathetic communication can be learned.
3. Successful clinicians rely on empathetic communication and behaviors.
4. Clinician behaviors have an impact on the emotional state of patients.
5. Improving the ability to detect and empathetically respond to a patient’s emotional state improves clinical interactions.

In particular, by focusing on simulated clinical situations with patients experiencing or recovering from SUDs, Medship targets the specific lapse in medical training aforementioned. As a web application, Medship combines educational resources specific to SUDs with affective coachware designed to provide tangible feedback on expressions and contextually-informed responses. Furthermore, the logic governing the feedback of expressions in Medship is informed by real-world patients with a current SUD or a history of SUDs, through data collected in our multi-phase collaborative human clinical trial with Weill Cornell Medical College. Having such patient-informed data is critical, as expressions, reactions and engagement in clinical scenarios with patients with SUDs are highly contextually-dependent. This feature places Medship in a unique position among the market of affective computing based coaching technologies and coachware, as no other technology captures these medical nuances.

1.2 Background

1.2.1 Addiction and Stigma

Addiction is a treatable, chronic disease characterized by a loss of free will, with complex interactions between the environment, genetics, life experiences, and the brain regions associated with executive function, reward, memory, and stress [47][46].

Despite often harmful consequences, people with an addiction are often unable to stop substance use due to compulsive behaviors and substance interactions. Over time, these changes compound and dramatically increase an individual's risk, making it substantially more difficult to stop using substances. This counters the widespread belief, even among physicians, that individuals with a SUD are simply lacking in will power to stop their use.

Addiction is one of the most stigmatized social problems today, bringing extreme burden to the social lives of those experiencing a SUD. Stigma, a socially discrediting behavior, prohibits individuals with SUDs from proper care and recovery, even within hospitals. We know that prevention and treatment approaches are often highly successful, at the level of other chronic diseases. Yet, within the medical community, a long-standing, deep-rooted stigma toward patients with SUDs stands in the way of this treatment.

1.2.2 Physician Attitudes and Education

A primary goal of Medship is to bring awareness to clinicians about the importance of their interaction quality with patients with SUDs, through repetition and emphasis during the training. Empathy, the physician's ability to understand and relate to a patient's emotional state and communicate that, is highly related to the degree of stigma toward a patient with a SUD [17][27]. A physician's ability to communicate effectively with patients, which is a strong proxy for their ability to empathize with patients, is associated with better outcomes, through adherence to medication, fewer mistakes, and higher patient satisfaction [38]. However, the outward display of empathy decreases across medical training [23][25].

Physicians often hold deep biases toward individuals with SUDs, more than any other group of individuals with a medical or mental health diagnosis [4][13][14]. Physicians, psychiatrists, psychologists, nurses, and social workers have less regard for individuals with SUDs than they do toward individuals who suffered from other illnesses, such as diabetes and depression [19][40]. Further, these attitudes worsen over time. For example, residents just a few years post-graduation have significantly worse

attitudes toward patients with SUDs than recent graduates [14][4].

These negative attitudes dramatically worsen quality of care. Physicians who stigmatize patients with SUDs hold biases that the patients are manipulative and violent, among other negative traits [5][9]. Consequently, physicians express less empathy toward these patients, which is a critical aspect of medical care. As a result, the interactions between patients with SUDs and physicians deteriorate, causing poorer clinical outcomes [30][9]. However, it is not the case that all physicians hold deep-rooted negative biases toward these patients. Rather, a single group has much less: physicians who work with these patients [19][13]. This highlights the importance of providing medical school students and physicians time with and exposure to patients with SUDs.

1.2.3 Marginalized Populations

Marginalized populations are at a higher risk for developing SUDs. For example, over 20% of the LGBTQ+ community have a SUD, compared to 9% in the general population [42][2][22]. This is due to a variety of reasons, such as access to care, baseline discrimination, and structural inequalities [24][16]. These biases in healthcare are well known, and detract from the quality of care—resulting in shorter conversation length, poor interpersonal relations, and diminished patient recall of appointment content [26][16]. This disrupts the patient-physician relationship, causes anxiety, and worsens outcomes [3][37]. As a result, these individuals face compounded discrimination, and have some of the worst outcomes of any population [41][36] [35][39]. This further highlights the importance of providing medical trainees exposure to patients in these populations.

1.3 Thesis Overview

Central to the completion of this thesis was a multi-step, prospective, longitudinal study which was used to inform the design and creation of Medship. The aim of the study, which was a collaborative effort with Weill Cornell Medical College, was to

define the outwardly communicated affective states of physicians that were perceived most favorably by patients with SUDs. This was done by recording the clinical interactions between physicians and a group of diverse patients with SUDs, drawn from both majority and minority populations, including those in the LGBTQ+ and Black communities. Following this, additional individuals with SUDs observed the video and audio recordings and annotated timestamps with information on affective expressions and states.

The results of this study were used to create Medship, designed to improve physician awareness and management of affective states toward patients with SUDs. It is not enough for a physician to feel empathetic or think empathetically—they must also outwardly emote empathetically with their patients. The proposed technology addressed this head-on by providing real-time training feedback on expressive states. Embedded in the web-application are affective computing models aimed at determining, in real-time, the affective expression of the trainee, which is used as an estimate for how they might be perceived. The latter is important as it is the proxy for how empathetically the trainee is outwardly emoting, which drives the feedback that enables them to improve the affective signals they send to patients.

1.3.1 Part 1: Interviews

In a collaboration with Weill Cornell Medical College, we conducted an IRB-approved human clinical study to observe genuine interactions between physicians and patients with SUDs and then evaluate the aspects of these interactions that were viewed as most and least empathetic, from the perspective of individuals with SUDs.

The first phase of the study, detailed in Chapter 3.2, was conducted from January through March of 2022, involved collecting video and audio recordings of structured interviews between physicians and patients with SUDs. These interviews were structured such that the physician walked the patients through a series of questions related to their present substance use, history of addiction, and positive or negative experiences along the way. However, beyond this script, the interactions, which occurred through Zoom, were highly realistic and genuine to a modern patient-physician

telemedicine interaction.

1.3.2 Part 2: Raters

The second phase of the study, detailed in Chapter 3.3, which started in April 2022 and is ongoing as of July 2022, involved administering surveys to individuals who had or who were in recovery from a SUD. In these surveys, the participants were asked to view segments of the video and audio recordings from the clinical interactions from the first clinical trial and answer questions regarding affective expressions and rapport building: including if the physician is seen as favorable, what emotion they evoked in them and what emotion they believed the physician to be communicating. All of these were critical to helping us understand which expressions are most appropriate when dealing with patients who have a SUD, as well as which contexts they fit. This is the foundation on which Medship operates, as a science-backed coachware platform, informed by real-world patients with SUDs.

1.3.3 Medship

Finally, we created Medship, a web-application detailed in Chapter 4, where a trainee, such as a medical school student, interacts with a virtual patient in a clinically relevant situation. Medship is a virtual tool containing a source of resources related to SUDs and trauma-informed care, and an interactive experience designed to help physicians get feedback on how they might appear when interacting with patients with SUDs. Within Medship, there are two distinct training modules, each with a specific training goal. During each module, the trainee can use their webcam to be monitored by real-time affective computing models. Following the completion of the module, the trainee is given feedback on their interaction, informed by the results of our study. Further, within each module, the actors with whom the trainee interacts, change their emotional state. This allows medical school students and physicians to practice clinically-relevant situations related to SUD in an environment where there are no negative consequences.

1.3.4 Impact

Medship is a low-cost, web-application-based educational asset for use by both individual physicians and academic centers as a part of their addiction medicine curriculum. There are a total of 94,242 medical students currently enrolled in medical schools in the US, thereby offering a significant area to integrate the tool [1]. Given both the void of current affective training and lack of emphasis on SUD, and the significant number of potential trainees— this technology could reasonably help improve clinical outcomes for patients with a SUD across the U.S. with minimal cost. Current physicians and medical school students may both benefit from the virtual tool, which is first and foremost designed to improve patient outcomes.

To ensure this potential is met, the actual virtual tool must be reliable— which involves establishing acceptability and feasibility outcomes, including tool quality, system usability, and acceptability of the tool. Given the multi-institutional nature of this project, our collaborators at Weill Cornell Medical College will be taking this on as a next step. Acceptability will be assessed with various measures. We will use the System Usability Scale. A score of greater than 72.75 is considered the benchmark for “good” or satisfactory [6]. Regarding tool quality, we will use the Computer Tool Rating Scale. An average of 3 or greater is considered acceptable [29]. We will also examine general treatment acceptability, with a target benchmark of greater than 28 on the Abbreviated Acceptability Rating Profile [45]. Further, we will administer a questionnaire to participants in our pilot study of medical students, specifically: before the study, immediately following the study, and six months later. This includes the Medical Condition Regard Scale (MCRS) [10].

Chapter 2

Related Work

2.1 Virtual Training

Medship is designed to integrate goal-directed scenarios with real-time feedback, within an easy to use and fun environment. Similar work has shown that medical trainees prefer using technology that is fun, develops clinical skills, or is goal-directed [8][21]. Thus, in the design of Medship, we took care to incorporate each of these.

Furthermore, Medship offers personalized routes to cognitive and emotional improvement with interactions with patients with SUDs. Research has shown that there are a variety of benefits to training with this structure, such as real-time feedback [44][11] and a non-accusatory approach—providing a low-stakes environment for the physician to learn within [31].

Virtual training methods have been proven to help individuals restructure and relearn cognitive-emotional responses to certain situations. Such technologies have been helpful ranging from counseling for preconception care, chronic disease care, and mental health [7][28][12][43]. Applying standardized technology within addiction care can ensure reliable, valid and unbiased evaluation of trainees as well.

The most similar related work is a technology called MACH (My Automated Conversation coach), a coachware designed to help improve interviewing skills using affective computing [20][17][18][38]. Similarly to MACH, Medship interacts with live affective data to provide real-time, personalized feedback, which is used to train

attitudinal change robustly.

2.2 Coachware

Broadly, Medship can be considered a novel form of coachware. Coachware, although in its infancy, describes the emerging market of automated coaching technology, which is just beginning to be incorporated into a variety of industries— such as public speaking, improving emotional intelligence, executive coaching, personal coaching, and video coaching. However, Medship is set apart by its research and institutional backing, whereas the majority of publically available coachware is commercial. To name a few, Orai and Speeko are in public speaking coaching; Cogito and Nice are in emotional intelligence coaching; Sounding Board, CloverLeaf, and Get Mee are in executive coaching; Replika and Rocky are in personal coaching; and Poised, uSpeak, Sesh, and Yoodli are in video coaching. Most of these companies rely on inputs from audio or other formats, with Poised, uSpeak, Sesh, and Yoodli being exceptions using video inputs. Combined, these companies have tens of millions of USD in funding, and are rapidly expanding and defining the emerging coachware industry. Although Medship is not currently poised to enter the commercial coachware market, there is a clear demand for affective coachware with tangible expression feedback.

Chapter 3

Study Design

3.1 Description

Central to the completion of this thesis was a multi-step, prospective longitudinal study in collaboration with Weill Cornell Medical College. Although the study is still ongoing, with lingering data collection from the final phase, the majority of it has been completed and preliminary results are available. The purpose of the study was to identify the affective expressions seen as most favorable, from the perspective of patients with SUDs. This aligns with one of the central beliefs of Medship— that it is not enough to think and feel empathetically, one must also outwardly express empathy.

Before beginning the study, we had to submit and get IRB approval for our study design protocol, at both MIT and Weill Cornell Medical College. Due to delays in processing, this took from August 2021 to January 2022.

The first phase of the study, conducted from January through March of 2022, involved collecting video and audio recordings of structured interviews between physicians and patients with SUDs. The second phase of the study, conducted from April through August of 2022, and still ongoing, involved administering surveys to individuals who had or who were in recovery from a SUD. Both of these were critical to helping us understand which expressions are most appropriate when dealing with patients who have a SUD, as well as which contexts they fit.

3.2 Part 1: Interviews

The structured interviews between physicians and patients occurred over Zoom, with myself or a collaborator on-call to facilitate the introduction and video capture. There were a total of 19 interviews conducted, between 5 patients and 10 physicians. A 20th interview was scheduled, but the patient did not show, and thus we cancelled it. The interviews were primarily run by the physician on-call and divided into a variety of tasks: setup, introductions, alcohol use, drug abuse, recovery, and experiences in the healthcare system.

In recruiting both the physicians and the patients for the study, we took care to ensure that both Group 1 (physicians) and Group 2 (patients) were diverse and representative. A full demographics breakdown can be found in the following figures in the appendix: A-1, A-2, A-3, A-4, A-5, A-6, A-7.

3.2.1 Setup

Upon entering the call, participants were asked to do the following:

- Find a quiet and safe environment.
- Ensure your computer is plugged into a charging source.
- Try to find a plain background.
- Ensure that there is light originating from in front of your computer (such that your face is well lit on the Zoom video feed).
- Situate yourself in front of the camera so that no more than 5 inches below your collarbones are visible in the frame.

This was in order to ensure that the video quality was high, for use by the raters in part 2, as well as potential use by affective computing algorithms. At this point, the facilitating research turns off and hides their microphone and video, allowing the call to be just between the physician and patient. From this point onward, the call is run by the physician, who follow a series of defined tasks.

3.2.2 Task 1

The first task is a simple introduction, in which the physician is asked to greet the patient and let them know that they are about to perform a substance use assessment that involves questions about their experiences with substances, their recovery, and their experiences with the healthcare system.

3.2.3 Task 2

Moving forward, the physician carries out a comprehensive alcohol screening called AUDIT: Alcohol Use Disorder Identification Test. AUDIT contains the following questions:

- How often do you have a drink containing alcohol?
- How many drinks containing alcohol do you have on a typical day when you are drinking?
- How often do you have 5 or more drinks on one occasion?
- How often during the last year have you found that you were not able to stop drinking once you had started?
- How often during the last year have you failed to do what was normally expected of you because of drinking?
- How often during the last year have you needed a first drink in the morning to get yourself going after a heavy drinking session?
- How often during the last year have you had a feeling of guilt or remorse after drinking?
- How often during the last year have you been unable to remember what happened the night before because you had been drinking?
- Have you or someone else been injured as a result of your drinking?

- Has a relative, a friend, a doctor, or another health worker been concerned about your drinking or suggested you cut down?

3.2.4 Task 3

When the AUDIT is complete, the physician moves on to a drug screening called DAST: Drug Abuse Screening Test. Prefacing DAST, the physician indicates the following, for clarity:

"The various classes of drugs may include cannabis (e.g., marijuana, hash), solvents, tranquilizers (e.g., Valium), barbiturates, cocaine, stimulants (e.g., speed), hallucinogens (e.g., LSD) or narcotics (e.g., heroin, or other opioids). Remember that the questions do not include alcohol or tobacco"

Then, DAST contains the following questions:

- Have you used drugs other than those required for medical reasons?
- Do you abuse more than one drug at a time?
- Are you always able to stop using drugs when you want to?
- Have you had blackouts or flashbacks as a result of drug use?
- Do you ever feel bad or guilty about your drug use?
- Does your spouse or parents ever complain about your involvement with drugs?
- Have you neglected your family because of your use of drugs?
- Have you engaged in illegal activities in order to obtain drugs?
- Have you ever experiences withdrawal symptoms when you stopped taking drugs?
- Have you had medical problems as a result of your drug use?

3.2.5 Task 4

Next, the physician asks an open ended question related to the patient’s recovery journey:

“For the next portion of this study interview, I’d like to ask a bit about your recovery journey. What has your recovery been like for you? What has helped? What barriers have you encountered and how have you dealt with them?”

3.2.6 Task 5

Finally, the physician asks another open ended question related to the patient’s experience with the healthcare system:

“For the next portion of this study interview, I’d like to ask a bit about your specific experiences with the healthcare system. What have your experiences been like in treatment? What have you appreciated from your providers in supporting your treatment? What haven’t you found helpful from your providers as it relates to your recovery or therapeutic relationship?”

3.3 Part 2: Raters

Before conducting part 2 of the study, we had to first create five surveys utilizing clips from the video recordings in part 1. To create these clips, we had to manually annotate each of the 19 videos for the points in which the physician was most or least expressive. We chose to select according to this metric, as we are interested in the expressions which patients with SUDs view as most and least appropriate, given a context. Furthermore, modern emotion detection algorithms work best with expressions that are further from neutral. Thus, we clipped the 19 videos into 104 segments, applied a de-identification algorithm such that only the physician was visible while the audio of the patient was distorted, and then organized these segments into subsets of 2-4 by video. Then, we organized the subsets into surveys, such that each of five surveys contained six subsets. This was done so that the raters could get a feeling

for the context specific to a single interview, while also being able to rate multiple interviews in a single session.

Following the completion of the video clip segments of the survey, we organized a series of questions for each rater to answer. First, nonspecific demographic questions:

- What is your age?
- Which of the following describes you[r identity] best?
- Which of the following describes you[r race] best?
- Do you identify in the LGBTQ population?
- How many years have you had opioid use disorder? Please enter whole numbers.
(For example: 7)
- How many years have you been in recovery? Please enter whole numbers. (For example: 7)

The breakdown of answers to these for the first 9 raters can be found here: B.1, B.2, B.3, B.4, B.5, B.6, B.7. Then, for each of the 2-4 clips within a subset, the rater was asked the following questions:

- Please put yourself in the patient's shoes. I found the way the physician responded...
- What was most helpful/likeable about their facial expression?
- What was most unhelpful/unlikeable about their facial expression?

Next, upon rating each of the clips within a subset, the rater was asked the following question about the physician in them:

- Please put yourself in the patient's shoes and respond to the following four questions regarding the four clips from this doctor. You will then move onto the next section with a new doctor presented. 1 (disagree) to 5 (agree).

- I feel that the doctor appreciates me.
- I believe the doctor likes me.
- I am confident in the doctor’s ability to help me.
- The doctor and I trust one another.

Finally, we recruited and administered the survey to individuals with SUDs. Currently, 9 raters have completed the survey, and we expect to administer it to a total of 20 raters.

Chapter 4

System Design

4.1 Overview

Medship is a web-application compatible with major web browsers such as Chrome, Firefox, and Safari, created with Flask and hosted through Heroku. Heroku is a platform which enables the cloud operation of the app and facilitates easy deployment of updates. This was chosen as it enables us to easily push updates and stay secure, all while helping us avoid holding specific hardware. Given the multi-institutional, long-term, and collaborative nature of this project, we thought that maintaining a physical server would be excessive and challenging. Currently, upon directing to Medship's URL, new trainees are required to supply an email, create a username, and choose a password. Furthermore, we are currently planning to implement access-control via Weill Cornell Medical School's secure single-sign-on, which will keep Medship only to our target trainees within the medical school.

A primary consideration for keeping Medship desktop-based was the requirement of access to microphones and webcams. Although mobile devices often have these, we had to deal with the additional requirement of proper backgrounds and lighting, which often are not provided in settings in which trainees would be using a mobile device. Thus, we decided to restrict access to desktop computers and laptops which have microphones and webcams, and prime trainees to set up their background and environment to help our affective computing work optimally.

On the backend, Medship is built from HTML, CSS, and JavaScript, with a few Python scripts to run the affective computing and set up Flask, a package for Python-based web application development. The primary interface is built off of CoreUI, a free package for clean interfacing and design. Using CoreUI allowed us to focus less on creating various HTML and CSS objects and designs, and focus more on designing Medship to suit our needs and host our computation. All of these design choices were intended to simplify the web development aspect of this system, while easily supporting Python-based computation.

Structurally, Medship is composed of eight sections: authentication, landing, home, about, expression practice, training, background information, and references.

- Authentication - contains the Login and Register pages
- Landing - contains the phrase "How do your patients experience you?" (See Appendix A-8)
- Home - contains information about the project and collaborators (See Appendix A-9)
- About - contains information about the motivation of the project and the problems in care toward patients with SUDs (See Appendix A-10)
- Expression Practice - contains a game-like module to practice facial expressions (See Appendix A-11, A-12)
- Training - contains simulated clinical scenarios to practice outward expression of empathy toward patients with SUDs (See Appendix A-13, A-14, A-15)
- Background Information - contains information about SUDs, OUDs and caring for patients with addiction (See Appendix A-16)
- References - contained references for the cited information in texts across the web application

For most of these sections, the material is very straight-forward and easy to maneuver. However, the sections of Expression Practice and Training contain more sophisticated modules with a variety of nested pages.

4.2 Expression Practice

The first of the two major modules of Medship is Expression Practice, where trainees can practice expressing themselves visually with some minor feedback. The idea behind Expression Practice is that many people might not understand how they appear at certain times. It is one thing to think you are happy, angry, or empathetic; it is another thing to appear that you are one of these. Further, there are not great practice tools out there for helping you learn to make different expressions appropriate to a given context. Although the rise of web meetings like Zoom have given people a lot of time to look at themselves during meetings, for many, it is not intuitive to make an expression. Also, while listening intensely to a patient, it is easy to forget what one's own face might do, what messages it might send. Thus, we developed Expression Practice to help with this.

Within Expression Practice, there are two sections. The first is a landing page where trainees can select an expression they want to practice. For this, we simply embedded a few large buttons for ease of operation. When a trainee selects a button for an expression, they are taken to the second section: expression matching. In expression matching, for example, trainees see a certain expression that most people would label as "happy" on the right side of the page, and a video stream of themselves on the left side. The task here, is simply to match their face with the emoji, which can be surprisingly harder than it seems. Upon completing the expression matching, the trainees return to the landing page to select a new button.

Another purpose of this section is to give transparency for how well the system works or does not work on the individual trainee. A major limitation of this type of affective computing is that it may not work as well as we intend, based on a variety of factors including but not limited to individual differences, facial hair, clothing, acces-

sories and jewelry, background noise, and computer hardware. Thus, the expression matching section gives the trainees an opportunity to evaluate for themselves how well the models work on them. When they make an expression that they and others might perceive as happy, does the model predict that their expression is happy? If not, might this be due to simple fixes, such as lighting in the room or removing a hat? In the case that it simply does not work as intended, this trainee should be fully informed, and they can proceed or not with that in mind. If that is the case, Medship has been designed to still provide informative value to trainees, even in the absence of perfectly functioning models.

4.3 Training

The second major module of Medship is Training, where the primary study-informed logic will be housed. The idea behind Training is that trainees will select a clinical scenario with a simulated patient and practice contextually-relevant skills, specific to the patient. The general flow is that a trainee will choose a clinical scenario on the landing page, test their webcam setup, read some background information specific to the issues addressed within the scenario, read a short prefacing text to the scenario, and then engage in a series of tasks with a simulated patient in a variety of emotional states— such as asking a sad patient about their history of opioid use. These patient interactions are relatively short, from 10 seconds to a minute, and each focus on a particular skill or task interest. Further, the interactions are set up to feel like a telemedicine visit with the patient, where the trainee can see the camera of both themselves and the patient. Finally, during the interactions, the trainee’s webcam will be recorded and analyzed using affective computing software running in the Python back end. This analyzed recording is shown immediately after the interaction, alongside "Dr. Superstar", a real-world expert physician, who this thesis committee believes has responded appropriately. The trainee is then able to compare themselves against Dr. Superstar in a segment we like to call Expression Karaoke, where the closer one is to Dr. Superstar, the better they do.

In the current proof of concept system, there is only one Dr. Superstar with whom the trainees can compare. This is an obvious limitation of Medship, as there are significant differences among people, such as the culture in which they were raised, that impacts their experiences and affective expression. Ideally, this would be accounted for by having multiple Dr. Superstars, giving the trainee an opportunity to directly compare themselves to someone who is more like them. Further, the trainee might choose to see the same reaction from the perspective of multiple Dr. Superstars, which might allow them to learn different strategies for improving their interactions. This would allow cultural nuances to be more appropriately displayed and embedded in the training, and might improve the quality of Medship.

This Training flow involved a variety of important technical design decisions. Namely, how to display the virtual patient, which affective computing to use, and how to provide feedback to the trainee. With respect to the virtual patient, we had the options of creating a cartoon, developing a more sophisticated and lifelike avatar, or hiring actors to portray a part. Originally, the plan was to create or purchase a lifelike avatar, but due to budgetary and time constraints, we ultimately resorted to hiring actors. One benefit that actors provide, is that we can have them very easily portray a variety of expressions and visual states, which computationally would be extremely challenging.

With respect to choosing an affective computing model, we originally intended to train our own models given the data from our trials. However, due to IRB timeline delays, this was not possible. Thus, we resorted to two expression detection models: haarcascades and OpenFace. Haarcascades is a simple, Python and OpenCV based model that we use to do real-time face detection and classification of simple states commonly perceived by people as being happy, sad, angry, disgusted, and neutral. Primarily, this is used within Medship as a fun way to ensure trainees set up their camera and background correctly, such that they can be seen and classified by a simple model. Furthermore, it is used for Expression Practice. However, Haarcascades is not elegant enough to provide substantially meaningful data, only outputting bounding boxes around detected states with a confidence score. Thus, for the Training we

use OpenFace, a much more sophisticated Python package for expression detection. OpenFace allows us to do more comprehensive facial action unit detection and run on videos, collected from trainee cameras. Although OpenFace gives us more reliable results and more data, it is similarly not without flaws.

Modern affective computing and machine learning models for expression detection have a variety of limitations that hinder performance. Namely, they are trained on publically available image datasets, which are often aligned with standard US population demographics. A consequence of this is that many models have seen fewer images of individuals in minority groups during their training, which might result in worse task performance for individuals of those populations. Additionally, expression recognition models, such as haarcascades and OpenFace, are designed to predict the expressions that most people might perceive. Within this task, there is an inherent variance, which is reflected in the performance of the model. Expression detection models might not always provide results that most people find accurate. Finally, this accuracy is dependant on a variety of conditions, such as lighting, backgrounds, clothing, jewelry, accessories, and hair. Variations in these might cause the models to perform expression detection less robustly. To account for this, we have designed Medship such that even in the absence of accurate, robust expression detection, trainees might still find value and be able to learn. Furthermore, Expression Practice as a module is set up to inform trainees of how the models perform on them as an individual. That being said, Expression Practice is not a perfect tool. It is one thing to be accurately matched with an expression; it is another to be accurately matched during a longer task in which one might be deeply focused on an interaction.

With respect to providing feedback to the trainee, we aim to translate the findings from our clinical studies into Medship. Primarily, we have received data that provide a patient viewpoint on a variety of contextually-dependant clinical interactions with patients with SUDs. This dataset is able to provide the logic governing whether a response from a trainee is appropriate or not, as well as if they are building empathy and connecting with the patient in a meaningful way. However, given the timeline delays, we have only very recently collected this data. As a result, we have resorted

to designing a minimum viable product that provides feedback in the form of a comparison versus Dr. Superstar and facial detection. In the future, however, we aim to introduce the logic such that the trainee can receive more targeted, context-specific recommendations on how to improve their interactions. Also, more studies to collect and validate data will be needed before such recommendations can be generalized outside the currently scripted contents.

Currently, Training includes two clinical scenarios, and in the future we intend to host six. The first clinical scenario is focused on trauma-informed care, which involves five core principles: patient empowerment, choice, collaboration, safety, and trustworthiness. During this clinical scenario, trainees are presented with the following patient background information:

"You are working on the medicine inpatient floor. You are about to have an encounter with a 22-year-old female patient whom you have never seen before, but is here following an overdose from opioids. You read in her chart that she has no past history. The nurse also noted that the patient's best friend at the bedside commented that she has had a long history of trauma."

Following this, the trainee is asked to carry out an introduction, along with a variety of other typical introductory actions, such as asking the patient's name and pronouns, checking non-verbal cues, and inquiring about comfort. Then, the trainee works through a variety of tasks, such as asking them what brought them in, or asking about their history of opioid use. During each of these, the patient will respond in a different way, dependant on their mood, which can be something such as neutral, detached, or anxious. This allows the trainee to train a variety of clinically relevant tasks, each variations in patient mood. Once again, after each task, the trainee will be shown their response alongside Dr. Superstar and the output of the affective computing model.

The second clinical scenario is about meeting the patient where they are at. The focus of this session is to ask about the patient's life and practice motivational interviewing skills, to help the patient realize their actions are having negative consequences on their life. Primarily, this clinical scenario aims to train trainee to in-

tegrate the following harm reduction techniques: harm reduction is evidence-based and patient-centered, do not punish a patient for showing symptoms of a disease that they are seeking out support for, offer tools without trying to control their behavior, and understand that relapse is part of the recovery process. Before the interaction starts, trainees are presented with the following patient background information:

"You are working a shift in the Emergency Room. It's nine o'clock in the evening and you just saw a patient for a left forearm abscess that resulted from an untreated wound infection secondary to injecting drug use. Mr. Tamari is a 36 year old transgender man with a past medical history of focal segmental glomerulonephritis (managed well with corticosteroids) and hypertension (on lisinopril). He reports having a close family support system, which includes his wife of the past five years. He states he drinks socially and does not smoke any tobacco products. He has used heroin for the past 2 years on and off."

Once again, the trainee goes through similar tasks to the first clinical scenario, but this time is more focused on motivational interviewing. Similarly, the same structure of a task followed by a review with Dr. Superstar is present.

Combined, these two clinical scenarios offer a strong foundation upon which trainees can learn. They present a variety of common, important, and under-trained clinical experiences, with a variety of patient emotional states. This is hypothesized to be valuable to medical school students and physicians, as they do not get many opportunities to experience these situations, but behaving correctly within them is critically important to the health and stability of the patient.

Chapter 5

Results & Discussion

5.1 Clinical Study

Part 1 of our collaborative clinical study with Weill Cornell Medical College went smoothly. We were successfully able to recruit 10 physicians and 5 patients with SUDs, both with a variety of backgrounds and demographics. Further, we successfully conducted 19 of the 20 scheduled interviews, with the final interview being a no-show from one patient. Each of the 19 videos we recorded were of very high quality, with audible audio and high quality lighting and framing of the physicians. We were also able to successfully de-identify and cut each of the 104 clips, by removing the visuals of the patient and distorting the audio.

Part 2 of the study is progressing. Following a successful creation of five comprehensive surveys, as described in Section 3.3, each containing numerous clips of physicians interacting with patients with SUDs and a series of follow up questions, we were able to administer them to 9 of the 20 anticipated raters with a history of or active SUD. In response to the expressions of the physicians, the raters had varied opinions. Depending on the clip, raters found some physicians to appear disagreeable and unhelpful, empathetic and good at listening, or disinterested and annoyed. We found that 56.38% of the time, raters found the physicians to be helpful, 25.53% of the time neutral, 12.77% of the time unhelpful, and 5.32% of the time hurtful (See 5.1).

Response	#	%
Helpful	106	56.38
Neutral	48	25.53
Unhelpful	24	12.77
Hurtful	10	5.32

Table 5.1: Group 3 - Raters with SUD: Summary of the raters' perceptions of physicians' responses in clips from part 1

Survey	Concordance %	Cohen's Kappa
1	47.6	0.172
2	14.3	-0.041
3	60.0	0.309
4	27.3	0.200

Table 5.2: Group 3 - Raters with SUD: Metrics of inter-rater reliability for raters' perceptions of physicians' responses in clips from part 1

Further, in response to the same clips, there was concordance and discordance in opinion (See 5.3, 5.4). Given that we have received multiple responses for four of the five survey variations, we are able to compare raters (See B.8, B.9, B.10, B.11). We found that it was most common that both raters would find a physician helpful (29.76%) or one rater would find them helpful while the other neutral (30.95%). Interestingly, there was only concordance 35.71% of the time. Although the sample size is small, this supports the idea that individuals with SUDs may frequently interpret a physician's outward expressions of emotion differently. This finding supports the development of contextually-specific training for physicians, as well as an opportunity for Medship to show more than one "acceptable" way to respond, which the physician can adapt to their patient.

Upon inspection of the individual rater differences, it was apparent that each individual had significantly different perceptions of the clips (See 5.2). For Survey 2, only 14.3% of responses were concordant, compared to 47.6% for Survey 1, 60% for Survey 3, and 27.3% for Survey 4. When evaluating inter-rater reliability by Cohen's Kappa, we see that Survey 1 raters have a Kappa of 0.172 (slight agreement), Survey

	Helpful	Neutral	Unhelpful	Hurtful
Helpful	25			
Neutral	5	3		
Unhelpful	26	6	1	
Hurtful	15	2	0	1

Table 5.3: Group 3 - Raters with SUD: Agreement between raters on physicians' responses in clips from part 1

	Helpful	Neutral	Unhelpful	Hurtful
Helpful	29.76%			
Neutral	30.95%	3.57%		
Unhelpful	17.86%	7.14%	1.19%	
Hurtful	5.95%	2.38%	0.00%	1.19%

Table 5.4: Group 3 - Raters with SUD: Percentage agreement between raters on physicians' responses in clips from part 1

2 raters have a Kappa of -0.041 (poor agreement), Survey 3 raters have a Kappa of 0.309 (fair agreement), and Survey 4 raters have a Kappa of 0.2 (fair agreement).

Almost all of the "hurtful" ratings were from the perspective of a single rater (See B.11). This skew simply highlights the limitation of having such a low number of raters, given this phase of the trial is still pending further results. Individual differences of raters can have a dramatic impact on the data averages with such a low sample size. Inspecting rater justification, individual raters who found physicians to be hurtful often used words such as "pretending to care", "bored", or "annoying"; for helpful, they included words such as "smiled", "seemed interested", and "did not judge"; for neutral, they included words such as "nod" or "looked away"; and finally, for unhelpful, they included words such as "eyebrows" or "angry". Some of these, such as "seemed interested", are nebulous and not directly connected to a known set of facial action units. Nevertheless, it gives good insight into what aspects of an expression individuals pick up on when forming an opinion.

When asked about the physicians, from the perspective of the patients, more often than not, the raters somewhat or strongly agreed that the physicians were likeable,

	Strongly Disagree	Somewhat Disagree	Neither Agree nor Disagree	Somewhat Agree	Strongly Agree
I believe the doctor likes me.	2	3	14	15	20
I am confident in the doctor's ability to help me.	2	8	8	14	22
I feel that the doctor appreciates me.	2	3	13	15	21
The doctor and I trust one another	3	3	12	19	18

Table 5.5: Group 3 - Raters with SUD: Summary of the raters' perception of the physicians, taking the perspective of the patients

helpful, appreciative, and trustworthy (See 5.5, B.12, B.13, B.14, B.15, B.16).

5.2 Medship

Following the completion of this thesis, our collaborators at Weill Cornell Medical College aim to pilot Medship with a group of medical school students, pending IRB-approval. As a current project, we cannot assume the success of Medship as a tool used to improve the education of physicians and students just yet, as it has not been rigorously tested by the target group. Nevertheless, Medship, as a web-application hosting educational content, expression practice, and affective training specific to SUDs, has successfully been built as a first prototype and awaits use.

5.3 Discussion

Medical training specific to SUDs is substantially lacking, with physicians-in-training often never getting time to interact with actual patients suffering from SUDs, or even specific training targeting the outward expression of emotions and empathy. Medship,

as an educational resource aimed at helping physicians improve their outward display of empathy towards patients with SUDs, is targeting this issue and hopes to supplement current medical education. Currently, Medship's visions are nearly finished, pending the remaining results from the part 2 of our human clinical study. Upon the completion of that, we will be able to fully incorporate the viewpoint of patients with SUDs into Medship, finalizing the project and allowing us to begin testing. Following the completion of the IRB-pending pilot study at Weill Cornell Medical College, we aim to make any needed adaptations learned from the study, and if it is found helpful, to potentially deploy Medship to medical school students and physicians across the country.

Chapter 6

Future Work

6.1 Continuation of Medship

Due to previous and ongoing delays within the studies, Medship was reduced in scope. We originally intended to make the Training and Expression Practice modules much more sophisticated, and believe that they could be augmented in such a way for the future, given positive results from our IRB-pending pilot testing with Weill Cornell Medical College students. This would involve embedding the logic from our studies directly into custom affective models, such that each piece of feedback is truly patient-informed and aligns with our empirical results on the subject. This would solidify Medship as a science-backed training platform for current and future medical school students and physicians.

Furthermore, due to time constraints, we had to reduce the number of clinical scenarios that we could host. If the first two scenarios prove to be successful, a next step is to continue to create new ones, as there are countless clinical scenarios and diverse patient representations that would be useful for trainees. An ideal version of Medship is one that is comprehensive in training for the care of patients with SUDs.

Additionally, we would like to eventually record multiple Dr. Superstars, such that trainees could choose which Dr. Superstar they would like to be shown. This would allow for us to capture potential differences in culture that influence decision making and expressions, and convey that to the trainees.

Finally, we were unable to design a realistic virtual patient. It is unclear whether a virtual patient would be more successful than an actor as a training tool. One thing is certain, it would help Medship become more game-like, and possibly contribute to a more fun experience for the trainees. Nevertheless, it would be an interesting academic exercise to attempt an implementation of a virtual patient, and may contribute to the furthering of Medship as a tool and promoting its widespread use.

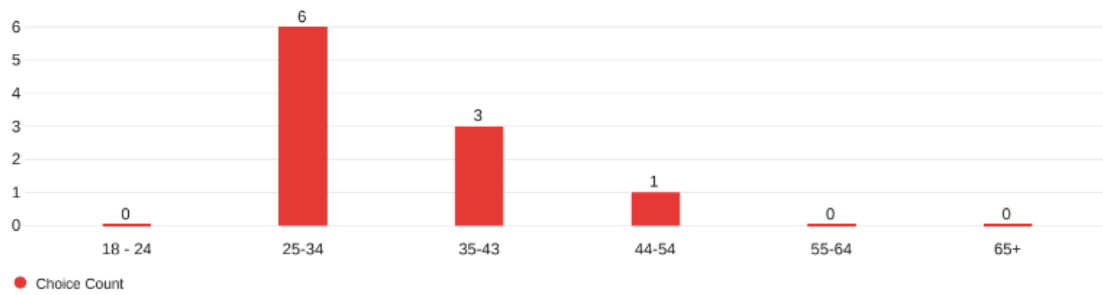
6.2 AI-Based Coachware

Medship, as previously discussed, is a new entry into the landscape of AI-based coachware technology. As an emerging market, coachware aims to revolutionize how we train a variety of different skills. Affective training is particularly suited to such a structure, as it is difficult to get quality, reliable training with traditional coaching. Most people learn how to control their outward expression of emotion through their childhood and upbringing. Only a few subgroups, such as actors, have affective training as adults. However, controlling one's expression is critically important to a variety of real-life scenarios, across industries and professions. Beyond the niche of medical interactions with patients with SUDs, products like Medship could transform how we view and train for person-to-person experiences. It is possible, and likely very easy, to take Medship as a model of affective training and apply it to a variety of use cases— such as customer service, sales, or interviewing. The primary hurdles are harnessing reliable, high-quality affective models and finding the use cases with the highest value. That being said, we believe that the model we have presented can and should be used as a template for future work, beyond the medical industry.

Appendix A

Figures

Q1 - What is your age?



Q2 - Which of the following best describes you?

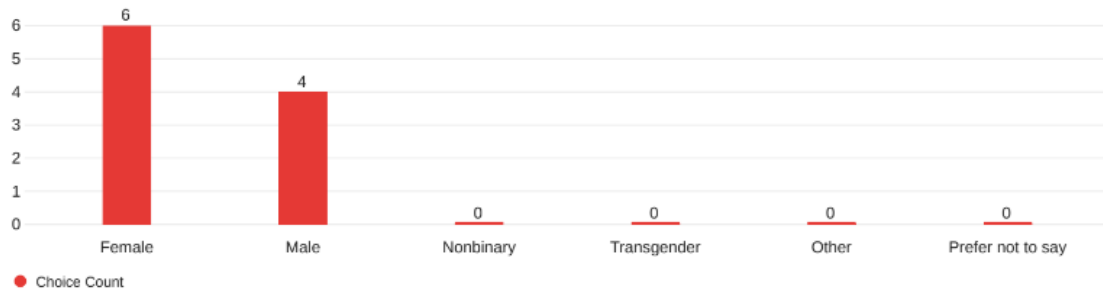


Figure A-1: Group 1 - Physicians: Demographics - Age and identity

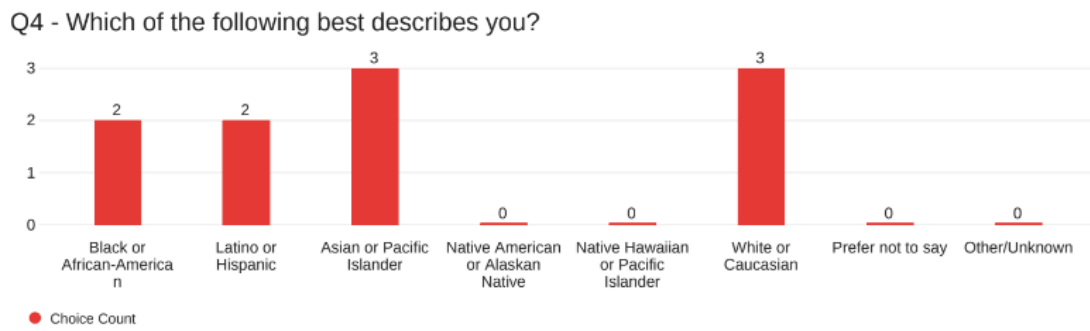
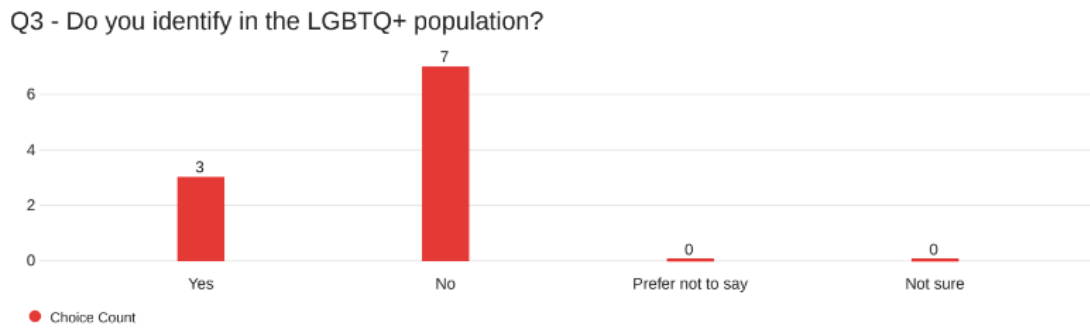
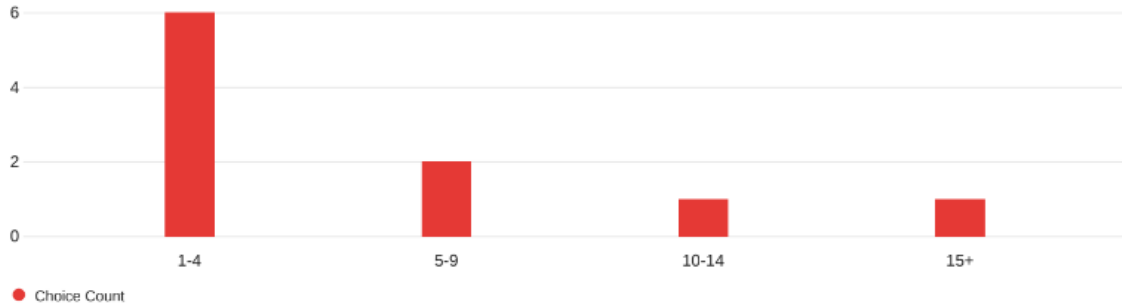


Figure A-2: Group 1 - Physicians: Demographics - LGBTQ+ and race

Q5 - How many years in practice are you?



Q6 - What is your specialty(s)?



Figure A-3: Group 1 - Physicians: Demographics - Years practicing and medical specialty

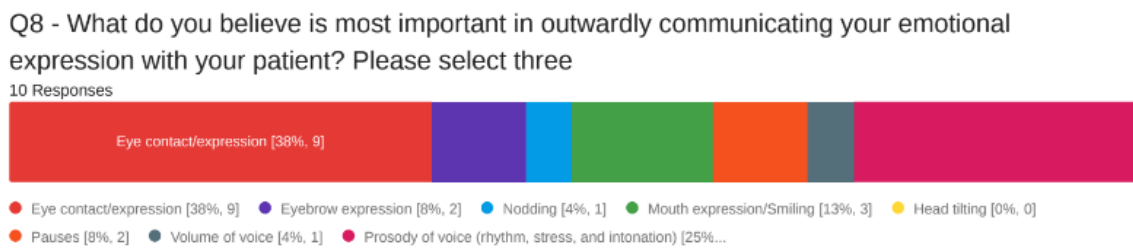
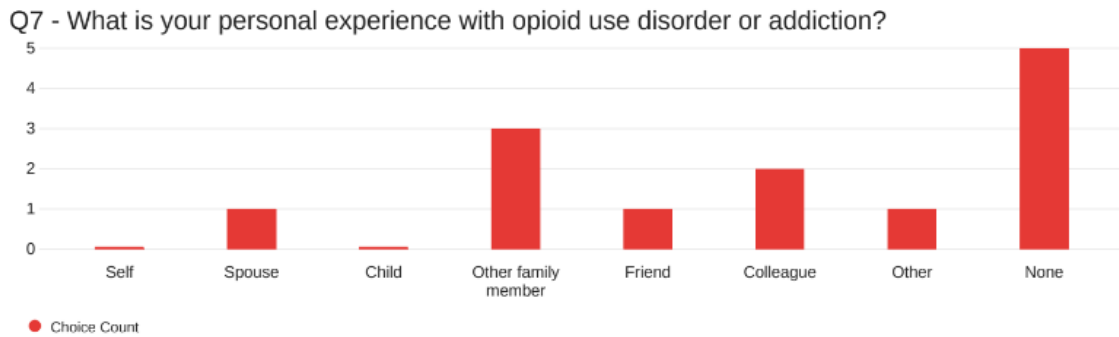
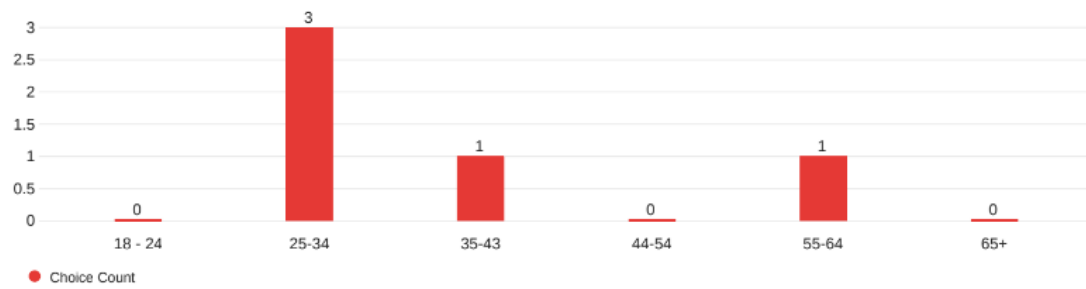


Figure A-4: Group 1 - Physicians: Personal experience with addiction and aspects deemed most important to see in a physician with respect to communicating emotional expression

Q1 - What is your age?



Q2 - Which of the following best describes you?

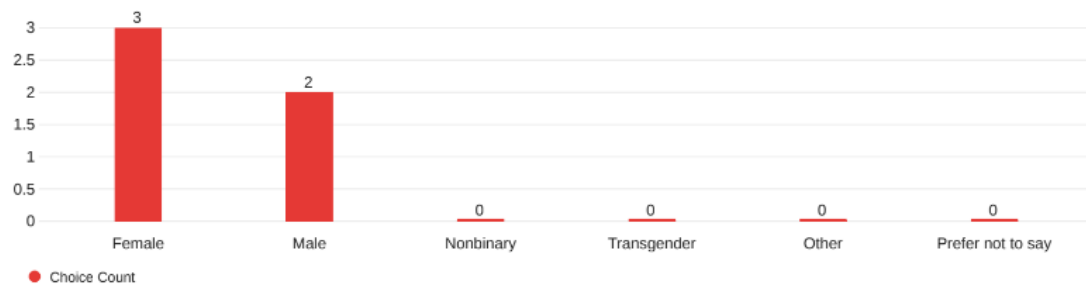
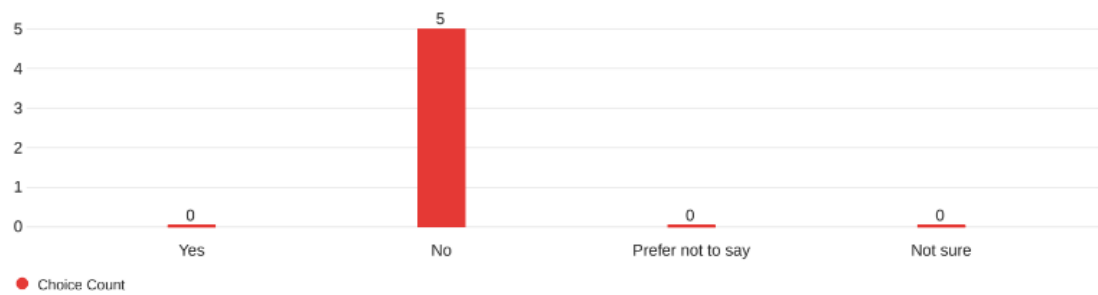


Figure A-5: Group 2 - Patients with SUD: Demographics - Age and identity

Q3 - Do you identify in the LGBTQ+ population?



Q4 - Which of the following best describes you?

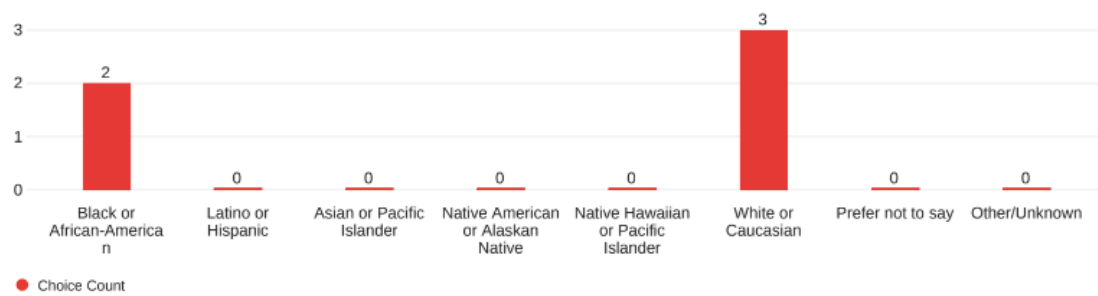


Figure A-6: Group 2 - Patients with SUD: Demographics - LGBTQ+ and race

Q8 - What is most important to see in your physician regarding the way they communicate their emotional expression? Please select three.

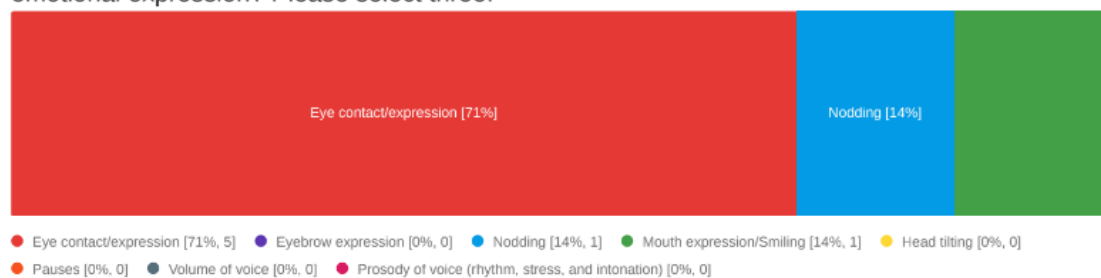


Figure A-7: Group 2 - Patients: Aspects deemed most important to see in a physician with respect to communicating emotional expression

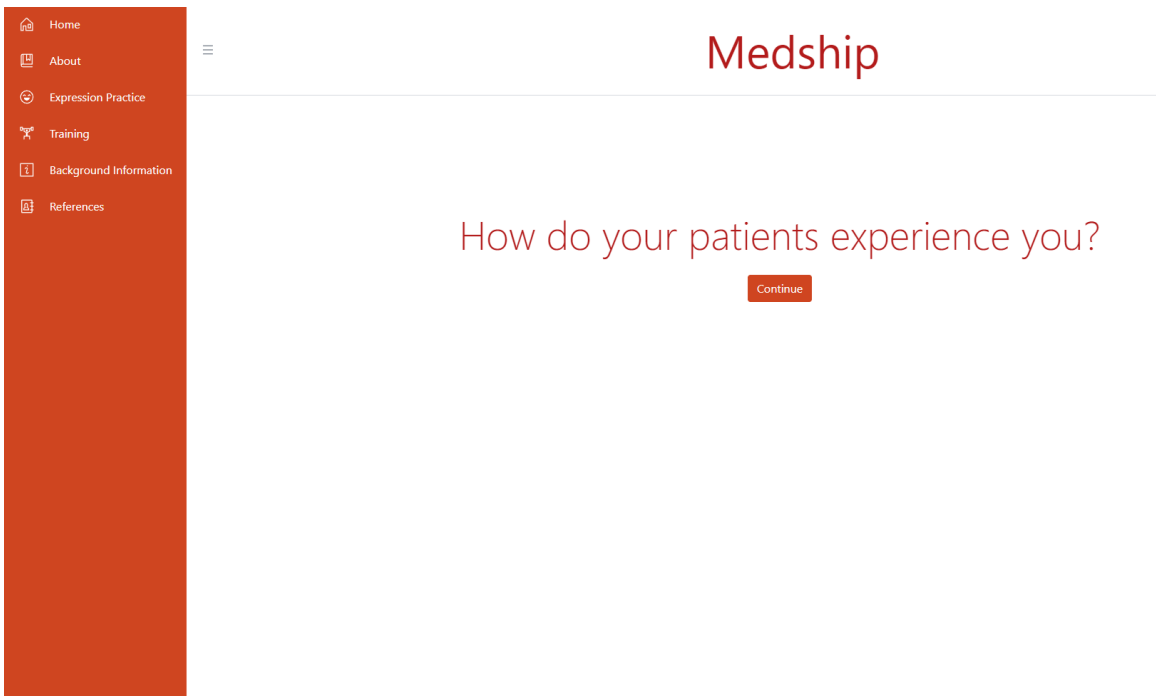


Figure A-8: Medship: Landing

About Medship

Medship is an initiative between researchers at Weill Cornell Medicine and MIT Media Lab. Our collaboration in Medship (like friendship) strives to create innovative ways for medical trainees to understand their own emotional experiences and best engage with the experiences of their patients, in ways that promote empathy, active listening, and compassionate dialogue.

Persons who inject drugs might face negative health outcomes due to stigma from the medical community, medical complications resulting from their use, and mutual mistrust with healthcare providers. Attempting to address these wrongs through compassionate and person-centered care is imperative for health justice, which we propose might be achieved through Medship – a greater understanding of our relationship with our patients.

Together, we are working toward a healthcare world that promotes the best therapeutic relationship between medical professionals and their patients.



Back

Continue

Figure A-9: Medship: Home

Introduction

Medical clinicians are trained to identify and treat diseases. Physicians, in the spirit of the Hippocratic Oath, vow to use their ability and judgment to help the sick, regardless of whom the sick might be. Thus, it is often assumed that the professionals in the medical community hold unbiased attitudes toward their patients, including toward individuals with opioid use disorder (OUD).

It turns out that the opposite is true. Physicians often have worse attitudes toward individuals with substance use disorders (SUDs), more so than toward individuals with other medical and mental health diagnoses. Further, these attitudes may deteriorate over time. Those who have been practicing for longer have been shown to have even worse attitudes toward patients with SUDs, and the worst toward opioid use disorder.

☰ **Stigma Towards Patients** ☰ **Clinician Attitudes** ☰ **Reasons for Negative Attitudes**

What is addiction stigma? ▼

What are the negative impacts of stigma? ▼

[Back](#)

[Continue](#)

Figure A-10: Medship: About

Expression Karaoke

Let's practice controlling your face and expression! We've found that people commonly have less control and understanding of their expression than they realize. So, we've designed **Expression Karaoke**, a tool to help you learn to control your expression.

Choose an expression to practice below! To skip **Expression Karaoke**, click "Continue".

Concerned Emoji

Happy Emoji

Back

Continue

Figure A-11: Medship: Expression practice landing

Expression Karaoke - Concerned Emoji

Match your face to the emoji! Click "Start Camera" to show your video, or "Start Hidden Camera" to record a video. When you are done, your analysis will be on the next page.

Start Camera Stop Camera



Back

Continue

Figure A-12: Medship: Expression practice example

Module Selection

Medship is composed of several modules, each one portraying a unique clinical vignette of interest, that focuses on specific and actionable training points. Click on the a tile below to train.

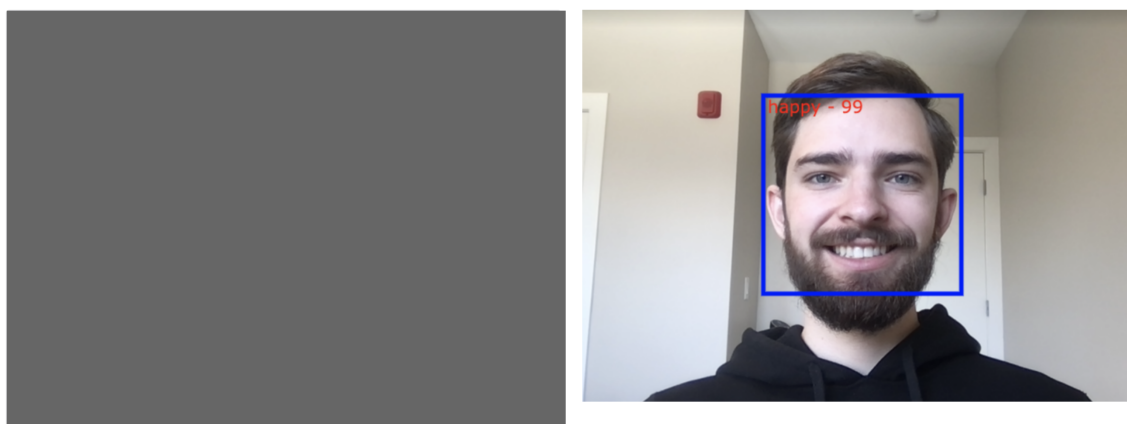
<p>Module 1</p> <p>Trauma-Informed Introduction</p> <p>Click here to proceed to the training.</p>	<p>Module 2</p> <p>Meeting the Patient Where They Are At</p> <p>Click here to proceed to the training.</p>	<p>Module 3</p> <p>Coming Soon</p> <p>This module is not available yet. Please check back later.</p>
<p>Module 4</p> <p>Coming Soon</p> <p>This module is not available yet. Please check back later.</p>	<p>Module 5</p> <p>Coming Soon</p> <p>This module is not available yet. Please check back later.</p>	<p>Module 6</p> <p>Coming Soon</p> <p>This module is not available yet. Please check back later.</p>

Figure A-13: Medship: Training landing

Video Setup

Our models work best when you position your face in the center of the frame, reduce background clutter, and keep bright lights out of frame. If a box identifies your face and expression, everything is working!

Start Camera Stop Camera



Back

Continue

Figure A-14: Medship: Training setup

Start Stop Next



Introduce yourself. Ask their name and pronouns.

Back

Finish

Figure A-15: Medship: Training example

Background Information

Diagnosing a Substance Use Disorder

According to The Diagnostic and Statistical Manual of Mental Disorders (DSM-5), a substance use disorder is a disease marked by a set of symptoms that include craving, tolerance, withdrawal, lack of control, and negative effects on personal and professional responsibilities. The DSM categorizes the disease by severity (mild, moderate, and severe) depending on the number of symptoms met. Individuals must meet at least two of eleven criteria to be diagnosed with a drug use disorder; these include the following:

- 1 Taking the substance in larger amounts or for longer than you're meant to.
- 2 Wanting to cut down or stop using the substance but not managing to.
- 3 Spending a lot of time getting, using, or recovering from use of the substance.
- 4 Cravings and urges to use the substance
- 5 Not managing to do what you should at work, home, or school because of substance use.
- 6 Continuing to use, even when it causes problems in relationships.
- 7 Giving up important social, occupational, or recreational activities because of substance use.
- 8 Using substances again and again, even when it puts you in danger
- 9 Continuing to use, even when you know you have a physical or psychological problem that could have been caused or made worse by the substance.
- 10 Developing tolerance: needing more of the substance to get the effect you want.
- 11 Development of withdrawal symptoms, which can be relieved by taking more of the substance.

Factors that Mediate Risk and Recovery

Recovery is marked by periods of remission and relapse. There is no simple cure for treatment. The disease is, however, treatable with continued efforts to manage symptoms, reduce the likelihood of relapse, and harm reduction. Ideal treatment includes a combination of psychosocial supports and medication management that can help people achieve sobriety.

Adverse Childhood Events (ACE): Increases Risk

Figure A-16: Medship: Background information

Appendix B

Tables

Age	#
18-24	1
25-34	4
35-43	1
44-54	1
55-64	1

Table B.1: Group 3 - Raters with SUD: Demographics - Age

Identity	#
Man	5
Woman	4
Non-Binary	0
Transgender	0
Prefer Not to Answer	0

Table B.2: Group 3 - Raters with SUD: Demographics - Identity

LGBTQ	#
Yes	0
No	9
Prefer Not to Answer	0

Table B.3: Group 3 - Raters with SUD: Demographics - LGBTQ+

Race	#
Black or African American	2
Latino or Hispanic	1
White or Caucasian	6
Asian or Pacific Islander	0
Native American or Alaskan Native	0
Native Hawaiian	0
Two or More	0
Other	0
Unknown	0
Prefer Not to Answer	0

Table B.4: Group 3 - Raters with SUD: Demographics - Race

Status	#
Recovery	7
Active	2

Table B.5: Group 3 - Raters with SUD: Demographics - Opioid use status

Years	#
0-10	1
10+	1

Table B.6: Group 3 - Raters with SUD: Demographics - Years with an active SUD

Years	#
0-9	5
10+	2

Table B.7: Group 3 - Raters with SUD: Demographics - Years in recovery from SUD

	Helpful	Neutral	Unhelpful	Hurtful
Helpful	8			
Neutral	6	1		
Unhelpful	4	1	1	
Hurtful	0	0	0	0

Table B.8: Group 3 - Raters with SUD: Survey 1 - Agreement between raters on physicians' responses in clips from part 1

	Helpful	Neutral	Unhelpful	Hurtful
Helpful	3			
Neutral	10	0		
Unhelpful	6	2	0	
Hurtful	0	0	0	0

Table B.9: Group 3 - Raters with SUD: Survey 2 - Agreement between raters on physicians' responses in clips from part 1

	Helpful	Neutral	Unhelpful	Hurtful
Helpful	9			
Neutral	6	1		
Unhelpful	2	0	2	
Hurtful	0	0	0	0

Table B.10: Group 3 - Raters with SUD: Survey 3 - Agreement between raters on physicians' responses in clips from part 1

	Helpful	Neutral	Unhelpful	Hurtful
Helpful	4			
Neutral	5	1		
Unhelpful	2	2	0	
Hurtful	5	1	1	1

Table B.11: Group 3 - Raters with SUD: Survey 4 - Agreement between raters on physicians' responses in clips from part 1

	Strongly Disagree	Somewhat Disagree	Neither Agree nor Disagree	Somewhat Agree	Strongly Agree
I believe the doctor likes me.	1	0	0	4	7
I am confident in the doctor's ability to help me.	1	1	0	2	8
I feel that the doctor appreciates me.	0	0	1	4	7
The doctor and I trust one another	1	0	2	3	5

Table B.12: Group 3 - Raters with SUD: Survey 1 - Raters' perception of the physicians, taking the perspective of the patients

	Strongly Disagree	Somewhat Disagree	Neither Agree nor Disagree	Somewhat Agree	Strongly Agree
I believe the doctor likes me.	0	1	4	3	4
I am confident in the doctor's ability to help me.	0	4	2	4	2
I feel that the doctor appreciates me.	0	2	4	2	4
The doctor and I trust one another	1	1	4	3	3

Table B.13: Group 3 - Raters with SUD: Survey 2 - Raters' perception of the physicians, taking the perspective of the patients

	Strongly Disagree	Somewhat Disagree	Neither Agree nor Disagree	Somewhat Agree	Strongly Agree
I believe the doctor likes me.	0	0	2	2	8
I am confident in the doctor's ability to help me.	0	0	1	2	9
I feel that the doctor appreciates me.	0	0	1	4	7
The doctor and I trust one another	0	0	1	5	6

Table B.14: Group 3 - Raters with SUD: Survey 3 - Raters' perception of the physicians, taking the perspective of the patients

	Strongly Disagree	Somewhat Disagree	Neither Agree nor Disagree	Somewhat Agree	Strongly Agree
I believe the doctor likes me.	1	2	5	4	0
I am confident in the doctor's ability to help me.	1	3	5	3	0
I feel that the doctor appreciates me.	2	1	6	2	1
The doctor and I trust one another	1	2	5	6	0

Table B.15: Group 3 - Raters with SUD: Survey 4 - Raters' perception of the physicians, taking the perspective of the patients

	Strongly Disagree	Somewhat Disagree	Neither Agree nor Disagree	Somewhat Agree	Strongly Agree
I believe the doctor likes me.	0	0	3	2	1
I am confident in the doctor's ability to help me.	0	0	0	3	3
I feel that the doctor appreciates me.	0	0	1	3	2
The doctor and I trust one another	0	0	0	2	4

Table B.16: Group 3 - Raters with SUD: Survey 5 - Raters' perception of the physicians, taking the perspective of the patients

Bibliography

- [1] AAMC. 2021 FACTS: Enrollment, Graduates, and MD-PhD Data. <https://www.aamc.org/data-reports/students-residents/interactive-data/2021-facts-enrollment-graduates-and-md-phd-data>, 2021.
- [2] M. Alexander. The new Jim Crow: Mass incarceration in the age of colorblindness. *The New Press*, 2010.
- [3] J. B. Asendorpf, R. Banse, and D Mucke. Double dissociation between implicit and explicit personality self-concept: The case of shy behavior. *Journal of Personality & Social Psychology*, 2002.
- [4] Jonathan Avery, Erin Zerbo, and Stephen Ross. Improving psychiatrists' attitudes towards individuals with psychotic disorders and co-occurring substance use disorders. *Academic Psychiatry*, 2016.
- [5] Bruce C. Ballon and Wayne Skinner. Attitude is a little thing that makes a big difference: reflection techniques for addiction psychiatry training. *Academic Psychiatry*, 2008.
- [6] Aaron Bangor, Philip T. Kortum, and James T. Miller. An empirical evaluation of the system usability scale. *International Journal of Human-Computer Interaction*, 2008.
- [7] Timothy Bickmore and Toni Giorgino. Health dialog systems for patients and consumers. *Journal of Biomedical Informatics*, 2006.
- [8] Stuart Brown. Play: How it shapes the brain, opens the imagination, and invigorates the soul. *Penguin*, 2009.
- [9] Noah A. Capurso and Daryl I. Shorter. Changing attitudes in graduate medical education: A commentary on Attitudes towards Substance Use and Schizophrenia by Avery et al. *The American Journal on Addictions*, 2017.
- [10] George W. Christison, Mark G. Haviland, and Matt L. Riggs. The medical condition regard scale: measuring reactions to diagnoses. *Academic Medicine*, 2002.
- [11] National Research Council. How people learn: Brain, mind, experience, and school: Expanded edition. *National Academies Press*, 1999.

- [12] David Devault and Ron Artstein. SimSensei Kiosk: A virtual human interviewer for healthcare decision support. *Proceedings of the 2014 International Conference on Autonomous Agents and Multi-agent Systems*, 2014.
- [13] Avery et al. Psychiatrists' attitudes toward individuals with substance use disorders and serious mental illness. *Journal of Dual Diagnosis*, 2013.
- [14] Avery et al. Changes in psychiatry residents' attitudes towards individuals with substance use disorders over the course of residency training. *The American Journal on Addictions*, 2017.
- [15] Avery et al. Improvement in residents' attitudes toward individuals with substance use disorders following an online training module on stigma. *HSS Journal*, 2019.
- [16] Avery et al. Is your own team against you? Implicit bias and interpersonal regard in criminal defense. *The Journal of Social Psychology*, 2020.
- [17] Bays et al. Interprofessional communication skills training for serious illness: evaluation of a small-group, simulated patient intervention. *Journal of Palliative Medicine*, 2014.
- [18] Bonvicini et al. Impact of communication training on physician expression of empathy in patient encounters. *Patient Education and Counseling*, 2009.
- [19] Gilchrist et al. Staff regard towards working with substance users: a european multi-centre study. *Addiction*, 2011.
- [20] Hoque et al. MACH: My automated conversation coach. *Conference: UbiComp 2013 - Proceedings of the 2013 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, 2013.
- [21] Kron et al. Medical student attitudes toward video games and related new media technologies in medical education. *BMC Medical Education*, 2010.
- [22] McCabe et al. The relationship between discrimination and substance use disorders among lesbian, gay, and bisexual adults in the united states. *American journal of public health*, 2010.
- [23] Neumann et al. Empathy decline and its reasons: a systematic review of studies with medical students and residents. *Academic Medicine*, 2011.
- [24] Pachankis et al. The burden of stigma on health and well-being: A taxonomy of concealment, course, disruptiveness, aesthetics, origin, and peril across 93 stigmas. *Personality and Social Psychology Bulletin*, 2017.
- [25] Patel et al. Curricula for empathy and compassion training in medical education: A systematic review. *PloS One*, 2019.

- [26] Penner et al. The effects of oncologist implicit racial bias in racially discordant oncology interactions. *Journal of Clinical Oncology*, 2011.
- [27] Pollak et al. Physician empathy and listening: associations with patient satisfaction and autonomy. *The Journal of the American Board of Family Medicine*, 2011.
- [28] Provoost et al. Embodied conversational agents in clinical psychology: a scoping review. *Journal of Medical Internet Research*, 2017.
- [29] Stoyanov et al. Development and validation of the user version of the Mobile Application Rating Scale (uMARS). *JMIR Mhelath Uhealth*, 2016.
- [30] Van Boekel et al. Stigma among health professionals towards patients with substance use disorders and its consequences for healthcare delivery: systematic review. *Drug and Alcohol Dependence*, 2013.
- [31] Williams et al. Racial/gender biases in student clinical decision-making: a mixed-method study of medical school attributes associated with lower incidence of biases. *Journal of General Internal Medicine*, 2018.
- [32] Wood et al. Physician education in addiction medicine. *JAMA*, 2016.
- [33] Center for Disease Control & Prevention. Drug overdose deaths in the u.s. top 100,000 annually. https://www.cdc.gov/nchs/pressroom/nchs_press_releases/2021/20211117.htm, 2021.
- [34] Center for Disease Control & Prevention. Understanding the epidemic. <https://www.cdc.gov/drugoverdose/epidemic/index.html>, 2021.
- [35] National Center for Health Statistics (US). Health, United States, 2015: With special feature on racial and ethnic health disparities. *Report No.: 2016-1232*, 2015.
- [36] Rachel R. Hardeman, Eduardo M. Medina, and Katy B. Kozhimannil. Structural racism and supporting black lives — the role of health professionals. *The New England Journal of Medicine*, 2016.
- [37] D. S. Jacoby-Senghor, S. Sinclair, and J. N. Shelton. A lesson in bias: The relationship between implicit racial bias and performance in pedagogical contexts. *Journal of Experimental Social Psychology*, 2016.
- [38] Bhautesh D. Jani, David N. Blane, and Stewart W. Mercer. The role of empathy in therapy and the physician-patient relationship. *Complementary Medicine Research*, 2012.
- [39] Christopher J. King and Yanique Redwood. The health care institution, population health and black lives. *Journal of the National Medical Association*, 2016.

- [40] S. C. Mahood. Medical education: Beware the hidden curriculum. *Canadian family physician*, 2011.
- [41] United States Department of Health & Human Services (HHS). National partnership for action to end health disparities: Offices of Minority Health. Washington, DC, 2011.
- [42] National Alliance on Mental Health (NAMI). LGBTQI, 2020.
- [43] Lazlo Ring, Timothy Bickmore, and Paola Pedrelli. An affectively aware virtual therapist for depression counseling. *Conference on Human Factors in Computing Systems (CHI) workshop on Computing and Mental Health*, 2016.
- [44] Diana Selmecky and Ian G. Dobbins. Metacognitive awareness and adaptive recognition biases. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 2013.
- [45] K. J. Tarnowski and S. J. Simonian. Assessing treatment acceptance: the abbreviated acceptability rating profile. *Journal of Behavior Therapy and Experimental Psychiatry*, 1992.
- [46] N. D. Volkow and R. D. Baler. Addiction science: Uncovering neurobiological complexity. *Neuropharmacology*, 2014.
- [47] N. D. Volkow and M. Boyle. Neuroscience of addiction: relevance to prevention and treatment. *American Journal of Psychiatry*, 2018.