2022

Smartphone-sourced data visualization in mental health

BOSTON UNIVERSITY

SCHOOL OF MEDICINE

Thesis

SMARTPHONE-SOURCED DATA VISUALIZATION IN MENTAL HEALTH

by

LUKE SANDERS SCHEUER

B.A., Bowdoin College, 2017

Submitted in partial fulfillment of the

requirements for the degree of

Master of Science

2022

Approved by

First Reader ____
Daniel Fulford, Ph.D.
Associate Professor, Department of Psychological and Brain Sciences

Second Reader____
John Torous, M.D., MBI.
Director, Digital Psychiatry Division, Beth Israel Deaconess Medical Center

SMARTPHONE-SOURCED DATA VISUALIZATION IN MENTAL HEALTH

LUKE SANDERS SCHEUER

ABSTRACT

Background. While smartphone digital phenotyping smartphone apps today can collect vast amounts of information on participants, less is known about how this data can be shared back with participants. Effective data visualization is critical to ensuring applications of digital signals are more informed, ethical, and impactful. But little is known about how sharing of this data, especially at different levels from raw data to analyzed data, impacts patients' perceptions.

Methods. We compared five different visualizations strategies, each a graph, generated from data created by the open source mindLAMP app, that reflected different ways to share data from simple amount of data captured to more complex clinical correlations. All graphs were shown to 28 participants during individual video interviews, and the graphs usability was measured via the System Usability Scale (SUS). Additionally, participants were asked about their comfort sharing different kinds of data, administered the Digital Working Alliance Inventory (D-WAI), and if they would want to use these visualizations with care providers.

Results. Of the five graphs shown to participants, the graph visualizing change in survey responses over the course of a week, received the highest usability score, with the graph showing multiple metrics changing over a week receiving the lowest usability score. Participants were significantly more likely to be willing to share geolocation data after viewing the graphs, and 25 of 28 participants agreed that they would like to use these graphs to communicate with their clinician.

Conclusions. Data visualization can help participants and patients understand digitally-sourced data and increase trust in how they are sampled and used to create visualizations. As data sourced from digital technology becomes more complex, simple visualizations may fail to capture their multiple dimensions and new interactive data visualizations may be necessary to help realize their full value.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

D-WAI	Digital Working Alliance Inventory
GPS	Global Positioning System
SAMHSHA	Substance Abuse and Mental Health Services Administration

INTRODUCTION

There is currently a critical need for increased access to quality mental health care in the United States. A 2021 study conducted by the Substance Abuse and Mental Health Services Administration (SAMHSA), found that approximately 53 million people, representing 21 percent of the adult US population were struggling with some form of mental illness, and 5.6 percent of the population, 14.2 million adults, had had some form of serious mental illness – meaning a disorder that significantly limited or altogether prevented major life activities such as depression, bipolar disorder, or schizophrenia – within the past year (Substance Abuse and Mental Health Services Administration, 2021). For these vulnerable populations, new modalities of care can help to increase both access to and quality of care. For example, due in no small part to the COVID-19 pandemic, virtual mental healthcare has been shown to be both feasible and effective, with 7.2 million, or just under 50%, of adults with severe forms of mental illness receiving some form of mental health services virtually in 2020 (Reay et al., 2020; Substance Abuse and Mental Health Services Administration, 2021). Thus, technology is already being used to great effect to increase access to and quality of mental health care.

However, in addition to using technology to facilitate access to clinician , the ubiquity of computers and smartphones allows for additional information that can be collected and used in a clinical setting. Digital devices can collect a variety of clinically useful data streams – and can even offer advantages over traditional pen-and-paper surveys or cognitive tests, such as remote data collection or trial-level reaction time data. The pandemic highlighted the utility of remote data collection and refocused attention on the use of apps to gather survey data, and has refocused attention on more advanced uses such as that of mobile technology to even measure of cognition (Singh et al., 2021).

Beyond what a desktop or laptop computer can provide, portable devices such as smartphones and other connected devices such as smartwatches yield a further data stream -- they are able to collect rich, real-time, and temporally dense data that can be used in a variety of settings. From simple longitudinal surveys to interactive assessments of reaction time, digital data from smartphones offers a new window in health (Lagan et al., 2021; Gansner et al., 2020; Henson et al., 2019).

Many of these newer digital signals, for example Global Positioning System (GPS) data, offer tremendous potential to bring new contextual data into mental health. For example GPS can be used to ascertain time spent at home or visiting new locations, or accelerometer and screen state data, which can be used together to infer sleep amount, quality, circadian routines or in conjunction with wearables like smart watches to understand fatigue are already ubiquitous or commonly used in the general population (Luo et al., 2020; Wisniewski et al., 2019). These data streams are often already collected by smartphones, which are owned by the majority of the general population. In addition, smartphone ownership rates among individuals with mental illnesses are not radically different from the population at large – meaning harnessing these novel data streams provides a powerful metric by which to gain additional insight into the mental status of patients – provided that clinicians are able to successfully introduce any apps used to their clients (NW et al., 2021; Iliescu et al., 2021; Torous et al., 2014; Mote & Fulford, 2019).

Smartphone applications exist that can and have been used to collect these digital data streams – indeed there has been a great deal of global adoption of one such app, the mindLAMP platform (Bilden & Torous, 2022; Vaidyam et al., 2022). Data such as that collected through apps like mindLAMP have a variety of clinical and diagnostic uses, from analyzing internet use habits in adolescent populations, to measuring how greenspace exposure is associated with the symptoms of individuals with schizophrenia (Gansner et al., 2022; Henson et al., 2020). During the COVID-19 pandemic, the mindLAMP app was used to collect digital phenotyping data in order to track the mental health status of 100 college students (Melcher et al., 2021).

But while it is easy to make apps available, there exist challenges to their uptake, which include participant concerns around sharing different forms of potentially sensitive data, a perceived lack of clinical value, and long-term issues with engagement. For example, the free app COVID-Coach from the Veteran's Administration retained less than 5% of users after 2 weeks (Jaworski et al., 2021). While there are many solutions to boost engagement, data visualization is an important avenue that to date has received less attention.

The challenge of engagement is not new. Older studies confirm that many mental health apps fail to keep users engaged long-term; usage rates of mental health apps drop to less than 5% within 10 days (Baumel et al., 2019; Possemato et al., 2016). While these studies do not offer a direct solution, each suggests that one significant contributing factor to lower engagement may be a failure to return data in a meaningful way. A 2020 review article by Polhemus et al. of the status of the visualization landscape found that apps that collect data without visualizing it for users are often found to be unengaging – but also that to date there is insufficient research into how to make engaging and usable visualization (Polhemus et al., 2020; Simblett et al., 2018). To increase user engagement and retention, then, it is necessary to investigate what makes a graph useful.

Another benefit of data sharing may be improved trust with digital health tools like smartphone apps. Many people are reluctant to share their digital signals, particularly those they consider sensitive; due to privacy concerns and failures to effectively communicate the reason or reasons for which data is collected (Parker et al., 2019). Data visualization offers a solution in that it can help people learn how their raw data is used, how that raw data can be transformed and presented in a privacy preserving manner, and how the analyzed forms of data relate to their health. Given the vast amounts of temporal data generated by digital devices and the early state of research therein, visualization is even more important as it offers a tool that may be more accessible and interpretable to patients than summary statistics.

A further advantage of data visualizations is that it can also provide a way to more effectively integrate measurement-based care into the mental health clinic. Measurement-based care refers to the process of regular administration of standardized questionnaires and surveys, like the 9-item Patient Health Questionnaire, a measure of depressive symptoms, or the 7-item Generalized Anxiety Disorder screener, which measures symptoms of anxiety (Kroenke & Spitzer, 2002; Spitzer et al., 2006). The results of these measures can then be used to inform care and track changes in a patient's mental status. While no measure alone can be independently diagnostic of mental illness, and therefore must be used in concert with a clinician, measurement based care can be a powerful tool for detecting patients who are struggling to improve or at risk for relapse, and there are ongoing growing calls to incorporate measurement-based care techniques into common use (Fortney et al., 2017; Lewis et al., 2018). Unfortunately despite evidence that measurement-based care improves patient outcomes, as well as the fact that patients have long seen the value in measurement-based care, only 20% of behavioral health providers currently incorporate it in their treatment (Dowrick et al., 2009; Lewis et al., 2018).

While there are barriers to introducing measurement-based care into any clinic, digital technology can help overcome them. One significant hurdle, for example, is the initial entry of data into an electronic medical record system; but a system or app where patients enter their own data through questionnaires eliminates this roadblock – and once the data is stored electronically, providers with digital systems that allow them to access measurement-based care tools more easily do use them more frequently (Steinfeld et al., 2016). Digital forms of data can provide value to both patients and providers – and data visualizations can make it easier to incorporate measurement-based care and improve clinical outcomes. Successful incorporation of digital data streams and the visualizations that can be made with digital data could facilitate the incorporation of measurement-based care into clinical settings.

Unfortunately, despite the potential and current uses of digitally-sourced data, current research on data visualization and its potential benefits on trust, engagement, and clinical usefulness remains sparse. Engagement research to date on mental health apps has identified a need to provide users with personalized content available across a range of devices – without it, risk of participant drop-out increases (Baumel et al., 2019; Quaedackers et al., 2020). Furthermore, data transparency is required in order to maximize willingness to share data – participants are much more willing to share data if they understand the purpose of their data; namely, that it is being used for their clinical care (Adanijo et al., 2021).

While the need for graphs that individuals with mental illness can use both on their own and in concert with physicians or other healthcare workers has been documented, most research studies focused on the design or implementation of one particular app or product instead of more generalizable knowledge that could be applied broadly (Polhemus et al., 2020). None to date have explored interactive visualizations, which may be particularly important for sharing complex temporal data gathered across numerous sensors. Interactive visualizations can show how survey measure scores change over time, give detailed insights into specific data points – for example showing specific survey responses on a day when a user's stress scores are high, or synthesize multiple data streams to yield information not available through any single measure – such as using time spent at home derived through GPS or accelerometer derived sleep quality and duration scores correlate with changes in mood or anxiety scores derived from measurement-based care measures. While numerous papers have examined engagement features of their own specific app and many call for good design and co-creation - few offer specific and generalizable principles (Bauer et al., 2018; Glomann et al., 2019).

Thus, this thesis focuses on three particularly relevant issues as informed by recent literature: how interactivity and data simplicity affect a graph's usability, how educating users about data affects their willingness to share it, and what graphs or visualizations users want to use by themselves versus which they want to use with clinicians or other mental healthcare providers.

In February 2022, we showed five graphs, already piloted in patient-facing studies in our lab and used by clinicians in our digital clinic, to 28 participants with previous experience using a digital mental health app (Rauseo-Ricupero et al., 2021). To investigate how data complexity, interactivity, and graph design affect the readability and usability of graphs, we administered the System

Usability Scale (SUS), a 10-item survey which measures the usability of a system (Bangor et al., 2008). To simulate interactivity, two graphs were shown alongside examples of ‘tooltips’ – an interactive window that, in a digital setting like one that would be used in a clinic, would allow a user to hover over data points to get additional information – here, data about specific survey responses, or explanations of different correlation values. To measure clinical usefulness and whether participants would want to use graphs in a healthcare setting, we measured the alliance between user and graph using the Digital Working Alliance Inventory (D-WAI), a measure of alliance between a user and a piece of software (Henson et al., 2019). To measure whether visualizations could change how comfortable users were with sharing different forms of data, we asked users to rate how comfortable they were sharing different kinds of commonly sampled data both before and after showing the graphs. Using the above methods, we explored how users can best understand and use their clinically relevant digitally-sourced data, the differences between static graphs versus more interactive visualizations utilizing tooltip hovering features, and how education and exposure changes a user’s level of comfort when sharing data.

METHODS

Section One - Participants & Recruitment

Participants for this study were recruited from a larger study investigating factors driving engagement with the mindLAMP app (Bilden & Torous, 2022; Vaidyam et al., 2022). Eligible participants for both this and the larger study were 18 years of age or older and reported at least moderate symptoms of stress as measured via the Perceived Stress Scale, but had not necessarily been diagnosed with a form of mental illness. Twenty-eight participants partook in a structured interview with Mr. Scheuer and completed three different measures during the study visit. The single study visit for each participant took on average 30 minutes. For specific participant demographic information, please see Table 1 below.

Table 1. Participant Demographics. Participant's self-reported demographic data is listed here.

	Percent of Total Population	Mean	SD
Gender			
Male	17.9		
Female	82.1		
Race/Ethnicity			
Asian	3.6		
Black or African American	10.7		
White	78.6		
Other	7.1		
Age		39.2	14.6

Section Two - Study Procedures

Prior to looking at any visualizations, participants were asked to rate how comfortable they were sharing five different forms of digital data collectable by a standard smartphone with their physician or care team: keylogging or content data from texts and emails (Keylogging), metadata concerning the number of texts or emails sent or received by a user (Metadata), Global Positioning System location data (GPS), accelerometer data (Accelerometer), or data from digitally administered surveys or questionnaires (Surveys), using a 5-point scale of

"Strongly Disagree", "Disagree", "Neither Agree nor Disagree", "Agree", or "Strongly Agree". This measure was repeated at the end of the visit as well.

Table 2. Comfort Sharing Data Measures. Participants were read the following statements one at a time and asked to indicate how much they agreed with them on a 5-item scale.

Statement	Abbreviation
I am comfortable sharing **Keylogging** data such as text or email content	Keylogging
I am comfortable sharing **Metadata** such as the number of texts or emails sent, but not the contents	Metadata
I am comfortable sharing **GPS** data, which measures my location	GPS
I am comfortable sharing **Accelerometer** data, which measures how and in what way I move	Accelerometer
I am comfortable sharing **Survey** data, which are my responses to surveys or questionnaires assigned by my clinician	Survey

Second, five graphs of gradually increasing complexity and analysis level were shown to the participants: a data quality graph (Data Quality), a graph showing the changes in two survey scores over a week (Survey Responses), a set of graphs showing, by day over the course of a week, time spent at home, step counts, and screen use data derived from analyzing passive data (Analyzed Passive), a summary graph showing weekly change in multiple metrics (Summary), and a correlation graph comparing multiple metrics (Correlation). Each graph was accompanied by a brief description (see Figures 1-5). For each graph, the participants took the System Usability Scale, a measure of usability which asks users to rate a series of ten statements pertaining to the ease-of-use of a technology on a scale of “Strongly Disagree”, “Disagree”, “Neither Agree nor Disagree”, “Agree”, or “Strongly Agree” (Bangor et al., 2008). Participants were also asked how often they would want to see or be shown these graphs if using them a clinical setting, from once a week to everyday, and were asked what device they would want them to be available on, either a computer, a phone, or both devices.

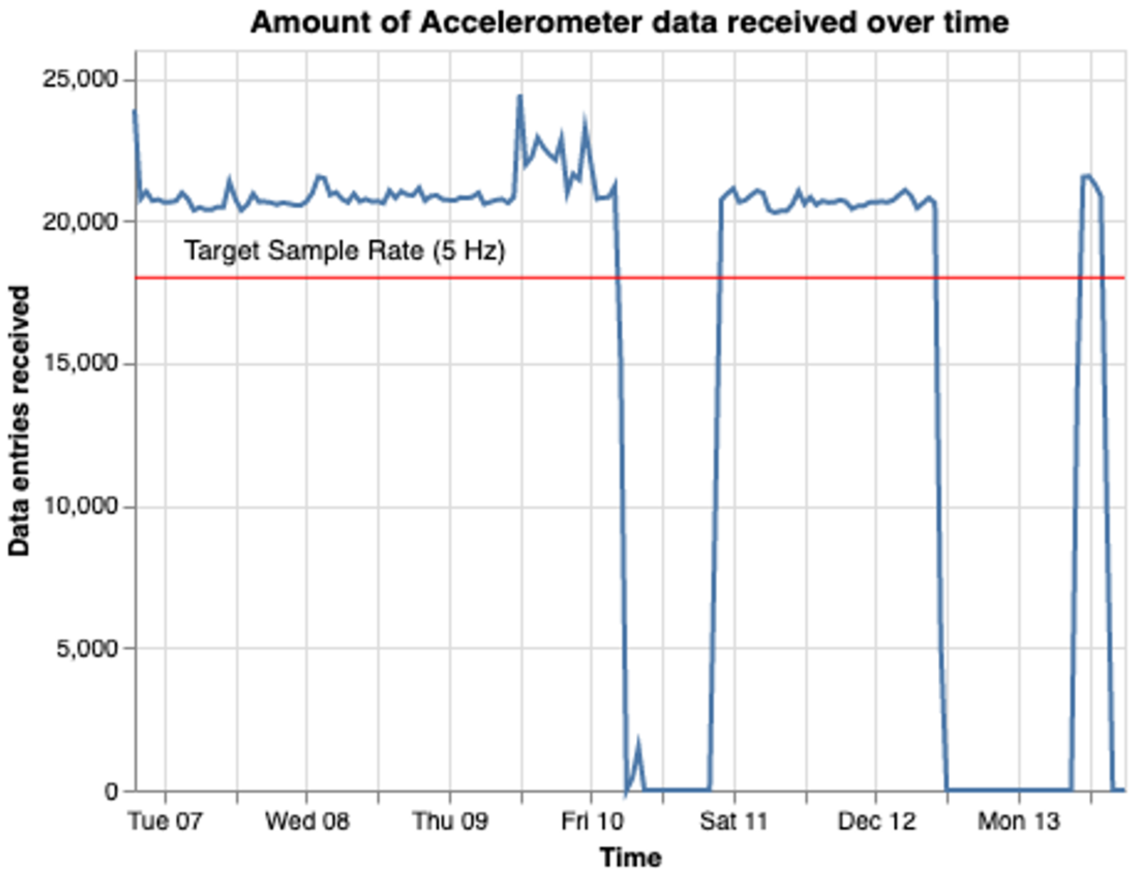

Figure 1. Data Quality Graph. This example of a graph plotting changing data quality over time was the first shown to participants. This text was read aloud when the graph was shown: “This first graph is a measurement of the amount of data (here, accelerometer) collected by your phone each hour over the course of a week. You could use this data to make sure that you were collecting enough data to get useful information from any analyses.”

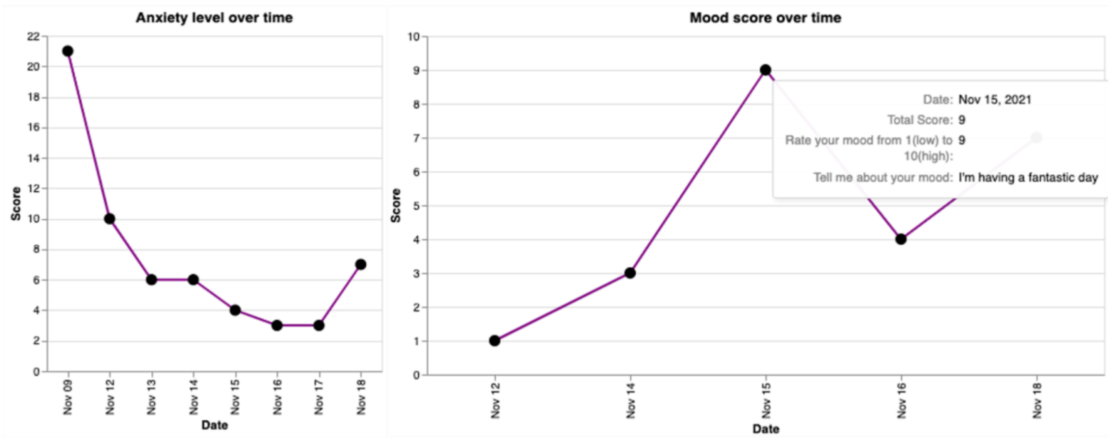

Figure 2. Survey Response Graphs. This example of survey response data was the second shown to participants. This graph contained an example of a tooltip, which can be seen in the right graph above. This tooltip contains the date the survey was taken, the score of the survey, and responses to each individual question. This text was read aloud when the graph was shown: “The second graph is an example of survey data collected over the course of around a week. You could see how your score changes from day to day. By hovering over a specific data point, you can see your specific responses to each question, as shown on the right.”

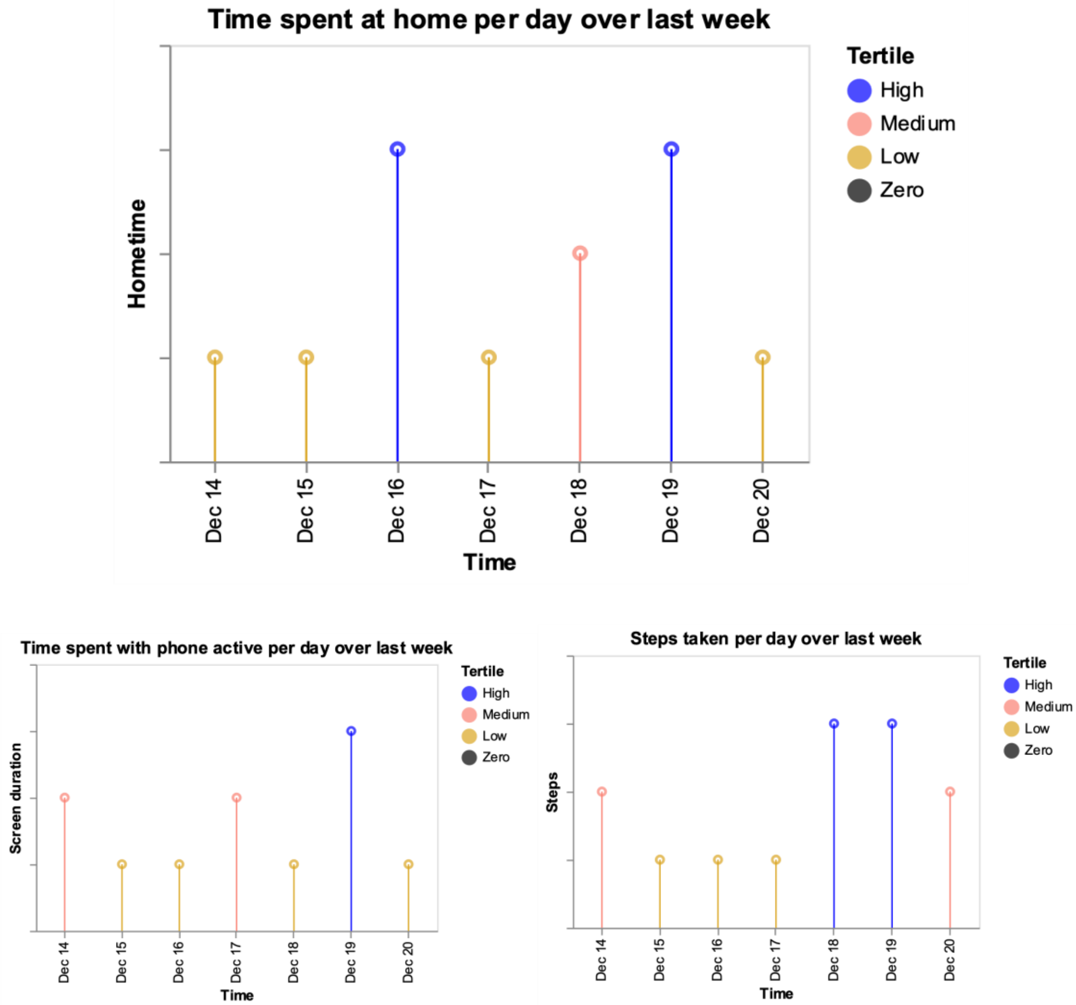

Figure 3. Analyzed Passive Graphs. These examples of data that can be sourced from passive data such as GPS or accelerometer readings were shown to participants third. The accompanying narration was: “These graphs are examples of data generated by analyzing data from sources such as GPS or an accelerometer. Here, you could see how your steps, time spent at home, and time spent using a phone have changed over the course of a week. Of note, all these scores are relative to themselves, not any standard measure – as such, when a measurement is “high,” that means it is high for you.”

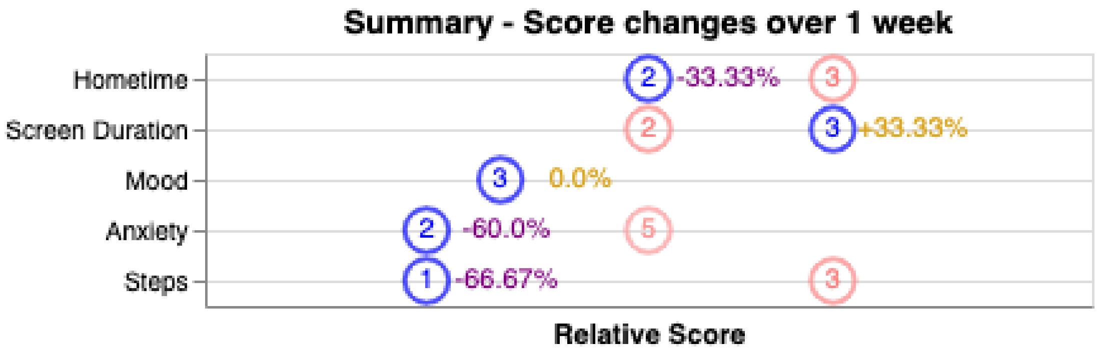

Figure 4. Summary Graph. This graph, which showed examples of relative changes in the various metrics used in prior graphs, was shown to participants fourth. The accompanying narration was: “This graph is a summary chart intended to show how multiple types of data change over the course of a week. Here, red points are “old”, and blue points are “new,” with the percentage change over a week, shown next to the newer point. For example, here hometime has decreased by about 1/3rd over a week. Mood is unchanged here.“

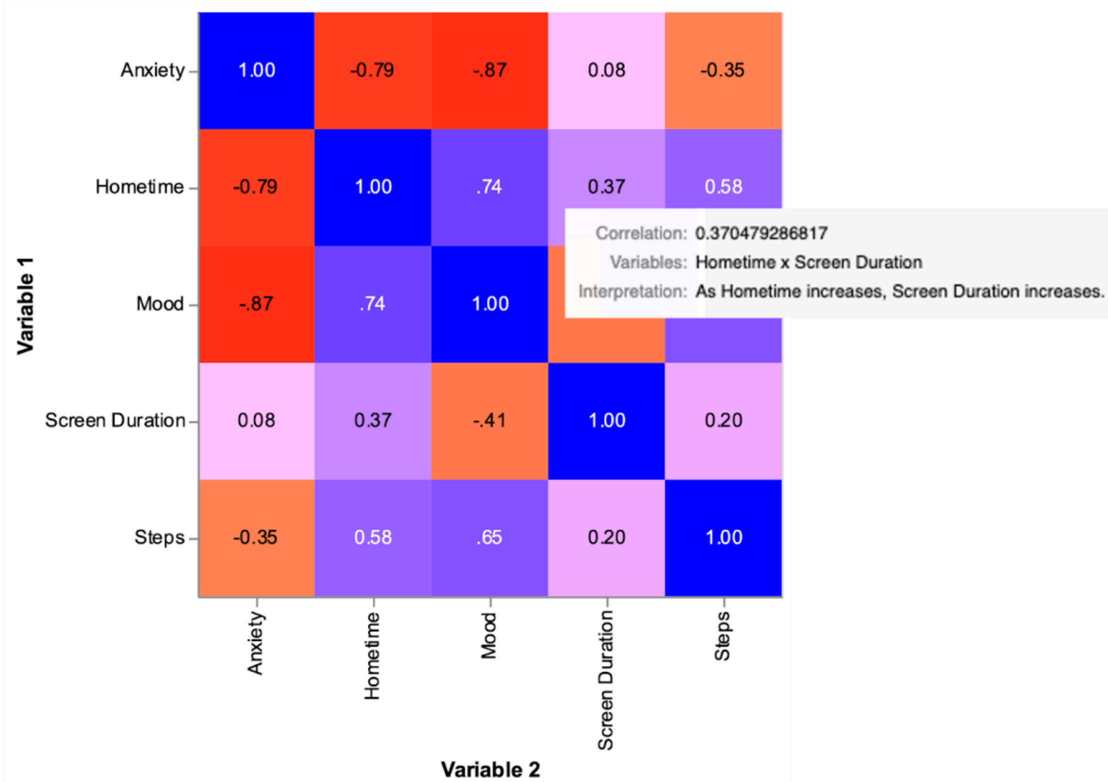

Figure 5. Correlation Graph. This graph, which shows examples of different correlations between the five example variables showed to participants thus far, was the fifth and final graph shown to participants. This graph, like the survey response graph, also contained a tooltip example, which reported the exact correlation value, the two variables being compared, and a brief interpretation of the correlation value, to increase usability. The text that was read aloud with this example was: “This graph is a correlation chart intended to show how some kinds of data change together. Higher correlations, meaning scores increase together, are shown in blue, with lower correlation, meaning scores change in opposite directions, are shown in red. For example, hometime and anxiety are negatively correlated – meaning as you spend more time at home, you are less

anxious. However, hometime and screen time are positively correlated – meaning more time spent at home is associated with more screen time. Like the second graph, hovering over a specific box would bring up a 'tooltip' with additional info; here, the exact value of the correlation, the two variables being compared, and a brief interpretation."

Table 3. System Usability Scale Items. All 10 items of the System Usability Scale are listed here. Items that are not reverse scored indicate higher usability if agreed with. Reverse scored items indicate lower usability if agreed with. Abbreviations are included for later visualizations.

Item	Reverse Scored	Abbreviation
I think that I would like to use this graph frequently.	No	Frequently
I find this graph unnecessarily complex.	Yes	Complex
I think this graph is easy to use or interpret	No	Easy
I think that I would need the support of a clinician to be able to use this graph.	Yes	Support
I found the various elements of this graph are well integrated.	No	Well Integrated
I thought there was too much inconsistency in this graph.	Yes	Inconsistency
I would imagine that most people would learn to read these graphs very quickly.	No	Quickly
I found the graph very cumbersome to use or understand	Yes	Cumbersome
I felt very confident reading this graph.	No	Confident
I needed to learn a lot of things before I could get going with this graph.	Yes	Learn

Finally, after looking at all the graphs, participants were instructed to imagine they were using all five graphs in a care setting, in concert with a clinician or clinical team, were read the six-item Digital Working Alliance Inventory (D-WAI) and asked to indicate how much they agreed with each statement on a scale of “Strongly Disagree”, “Disagree”, “Neither Agree nor Disagree”, “Agree”, or “Strongly Agree”. The D-WAI is measure of the working alliance between a user and digital method of care – it is a shortened and modified version of the Working Alliance Inventory used to measure the alliance between a patient and their clinician (Henson et al., 2019). Participants were also asked two additional items: first, if they felt the graphs could provide new insight about their mental health, and second if the graphs could help them communicate better with their clinicians.

Table 4. D-WAI and Added Questions. All 6 D-WAI statements asked are listed here. Asterisks indicate the additional items not present in the traditional D-WAI.

Statement
I trust these graphs to guide me towards my personal goals
I believe the graphs would help me to address my problem
The graphs encourage me to accomplish tasks and make progress
I agree that using the graphs are important for my goals
The graphs are easy to use and understand
The graphs support me to overcome challenges
These graphs give me a new way to look at my problems*
These graphs would help me communicate better with my clinician*

Section Three - Hypotheses

Based on both our previous research and evidence in the literature, we generated four hypotheses we hoped to test.

First, we theorized that the most usable, and thus most clinically useful, graph would be one that was neither too simple nor too complex; that is, that overly simple graphs like the Data Quality graph would not contain enough information to be helpful to a participant, while overly complex graphs like the Correlation graph would contain too much and thus be confusing. We tested this hypothesis by measuring graph usability using the SUS and comparing using an ANOVA test and post-hoc t-tests.

Second, we hypothesized that increasing understanding of data usage would lead to a corresponding increase in trust. To assess this, participants were asked about their comfort sharing types of passive data including GPS and keylogging data – and were shown graphs including measures derived from GPS (hometime), but not keylogging, after which their willingness to share data was assessed again. Comfort sharing GPS and keylogging data before and after seeing the graphs were compared to see if significant changes in comfort occurred using a paired t-test. Additionally, an ANOVA and post-hoc t-tests were conducted to see if comfort sharing any other forms of data changed significantly either before (“pre” condition) or after (“post” condition) the graphs were shown to participants.

Third, we hypothesized participants would think the graphs could be useful in a clinical setting, helpful to their goals, and provide value when working with their care team and clinician(s). To measure this, participants took the Digital Working Alliance Inventory to measure their alliance with the graphs, which we then compared to average D-WAI scores for popular apps in the mental health space. We also asked participants directly if the graphs would give them new perspectives on their data and if they would want to use the graphs with their clinicians.

Fourth, we believed that participants would want to have graphs available to them on both their mobile devices and computers, as opposed to exclusively one or the other – regardless of the graph itself. To measure this, we asked participants what devices they would want each graph to be available on and measured the results, then used an ANOVA test to compare desired device type across graphs.

RESULTS

Section One - Comfort with Sharing Different Forms of Digital Data

Comfort levels for each individual form of digital data were scored on a 0-4 scale: a score of 0 represents a participant indicating they strongly disagreed that they were comfortable sharing data while a score of 4 indicates they strongly agreed they were comfortable sharing data. By combining the score for all 5 types of data, total comfort sharing data was scored on a 0-20 scale. A one-way between subjects ANOVA analysis comparing comfort sharing different types data showed significant differences between kinds of data both before ($F(4,23)=12.98$, p-value$<.001$) and after($F(4,23)=17.98$, p-value$<.001$) viewing graphs.

Post-hoc analysis using t-tests revealed that, prior to seeing graphs, participants were statistically significantly less likely to feel comfortable sharing keylogging data compared to any other form of data (vs. Metadata: $p<.001$, vs. GPS: $p=.014$, vs. Accelerometer: $p<.001$, vs. Survey: $p<.001$); additionally, participants were significantly more likely to feel comfortable sharing survey or questionnaire data than any other kind of data (vs. Metadata: $p=.015$, vs. GPS: $p<.001$, vs. Accelerometer: $p=.045$). In addition, participants were significantly less likely to feel comfortable sharing GPS data than Metadata or Accelerometer data (vs. Metadata: $p=.004$, vs. Accelerometer: $p=.007$), and there was no significant difference between comfort sharing Accelerometer and Metadata.

After being presented with the five graphs, the comfort rating comparisons showed no noticeable changes: participants were significantly less comfortable sharing keylogging data relative to other data types (vs. Metadata: $p<.001$, vs. GPS: $p<.001$, vs. Accelerometer: $p<.001$, vs. Survey: $p<.001$); most comfortable sharing survey data (vs. Metadata: $p=.010$, vs. GPS: $p<.001$, vs. Accelerometer: $p=.032$); less likely to feel comfortable sharing GPS data than Metadata or Accelerometer data (vs. Metadata: $p=.009$, vs. Accelerometer: $p=.001$), and there was no difference between Accelerometer and Metadata.

However, there were some noteworthy changes in the pre and post comfort ratings. By way of a two-sided paired t-test, we observed a statistically significant increase in how comfortable participants felt with sharing GPS data ($p=.039$). Additionally, participants' comfort with sharing accelerometer data increased, but this change was not statistically significant ($p=.084$). We also found a decrease in participants' comfort sharing keylogging data – but this difference was also not statistically significant ($p=.211$). No other specific data types showed significant changes in comfort, and while overall comfort increased slightly, this change was also not statistically significant ($p=.18$).

Table 5. Comfort Score Means and Standard Deviations: Data showing the different averages and standard deviations for each type of data participants were asked about their comfort sharing. Each data type is described in more detail in the Methods section, except for Combined, which represents the total comfort score. Asterisks represent data types where comfort changed significantly after graph presentation.

Data Type	Mean (SD) Pre	Mean (SD) Post
Keylogging (0-4)	2.04 (1.52)	1.89 (1.72)
Metadata (0-4)	3.46 (.69)	3.5 (.58)
GPS (0-4)*****	2.75 (1.32)	3.1 (.99)
Accelerometer (0-4)	3.36 (.78)	3.57 (.57)
Survey (0-4)	3.82 (.39)	3.82 (.39)
Combined (0-20)	15.42 (3.46)	15.89 (3.31)

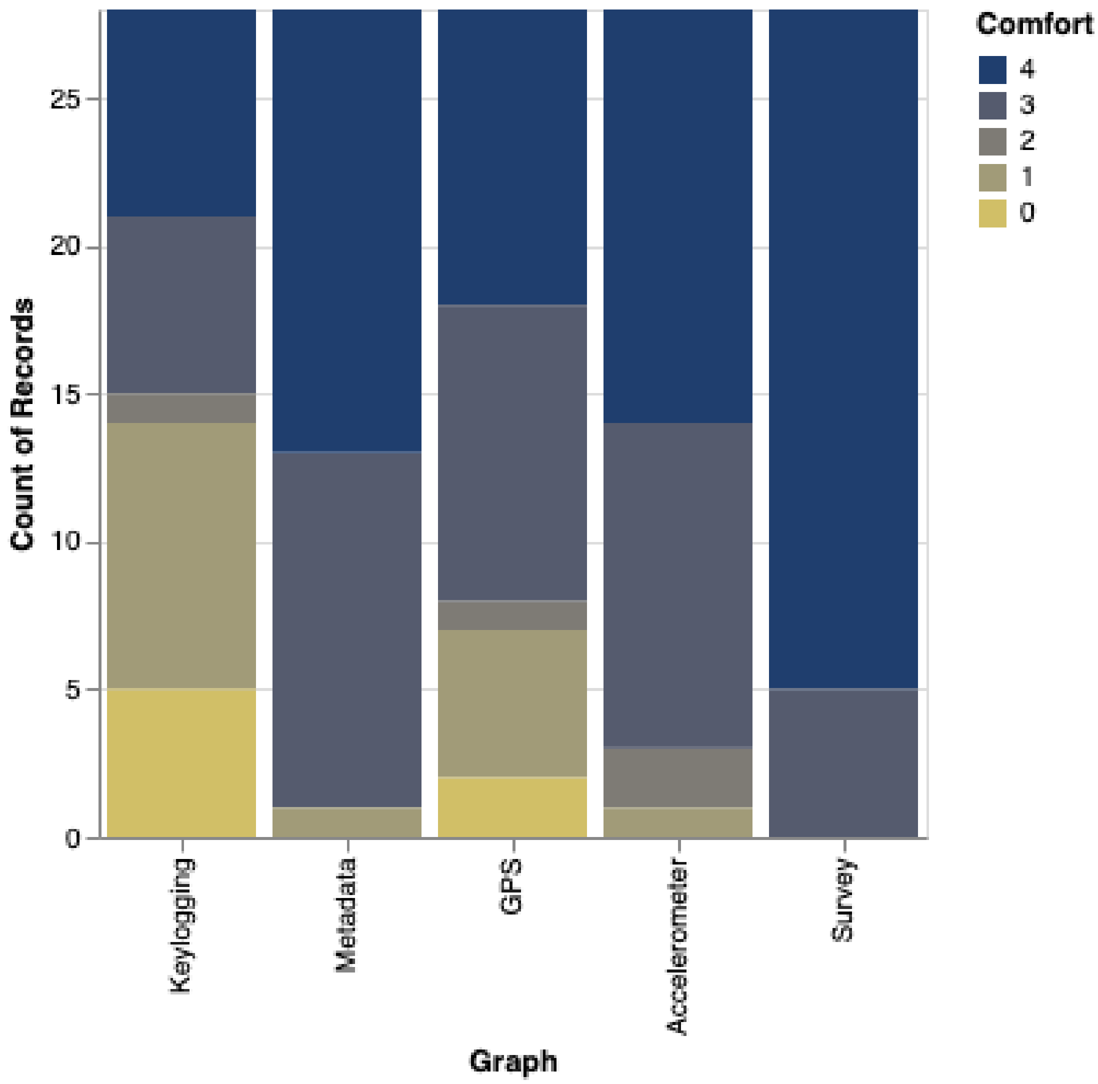

Figure 6, Part A. Stacked Bar Chart of Comfort Levels Before Seeing Graphs.

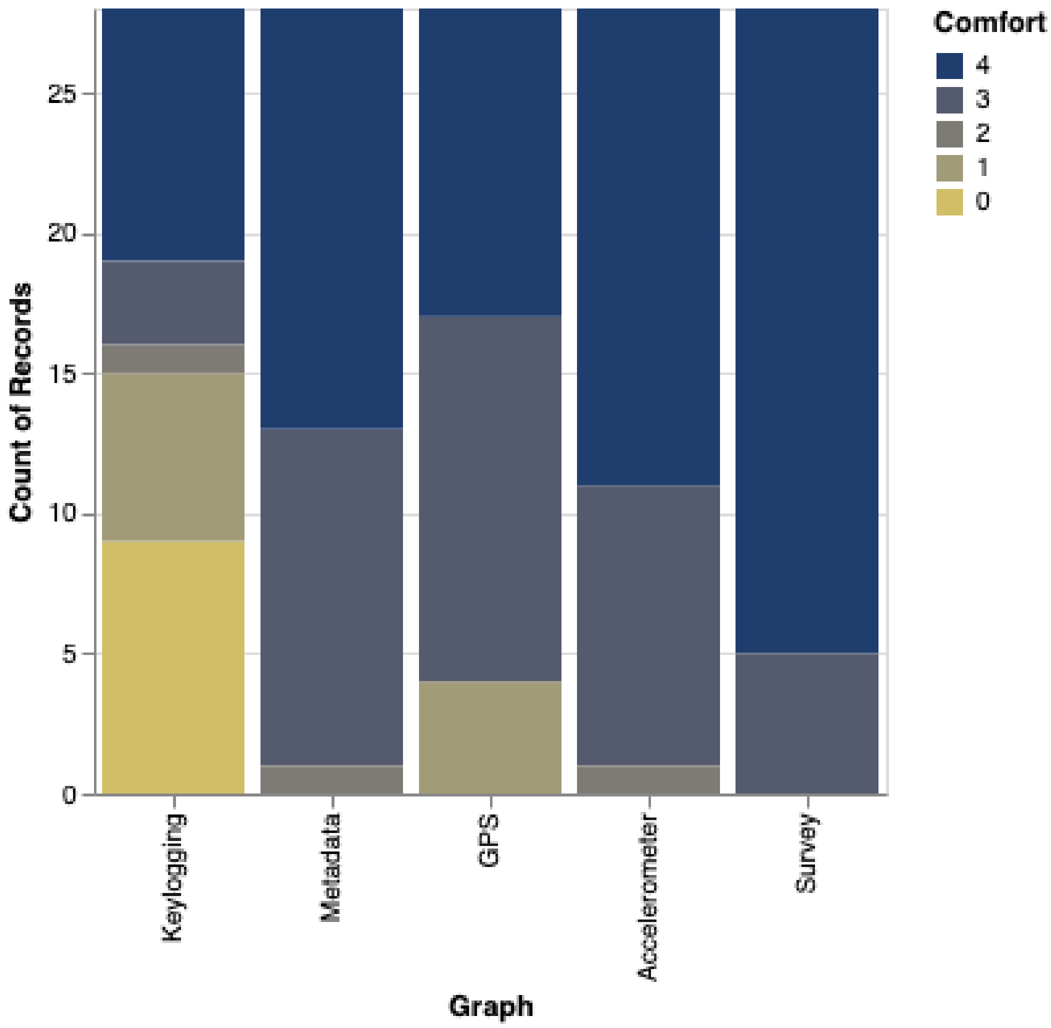

Figure 6, Part B. Stacked Bar Chart of Comfort Levels After Seeing Graphs.

These charts show the distribution of various participants' comfort levels sharing each type of data – Part A is before seeing graphs, Part B is after. A score of 0 indicates that a participant "strongly disagreed" they were comfortable sharing a specific type of data with physicians or their clinical team, while a score of 4 indicates they "strongly agreed" they were comfortable sharing that form of data.

Section Two – System Usability Scale

The 10 System Usability Scale items were each scored on a scale from 0 to 4, with a score of 0 corresponding to the lowest usability score for a given item, and a score of 4 representing the highest usability score possible for an item. Unlike the other two scales used, a score of 0 does not necessarily correspond to "Strongly Disagree" and 4 does not necessarily correspond to "Strongly Agree"; five items are reverse scored. The resulting 0-40 range was scaled by a factor of 2.5 to give a total range from 0-100. As SUS scores are not evenly distributed, we used a scoring system suggested by Bangor et al. to calculate an adjective rating, which is shown in table 4 (Bangor et al., 2009).

Table 6. Mean and Standard Deviation in SUS score, converted to Adjective Rating (Bangor et al., 2009) The mean System Usability Scale score for each graph shown, along with their standard deviation.

	Mean Score	Standard Deviation	Adjective Rating
Data Quality	69.29	23.44	OK (51-72)
Survey Responses	85.80	17.63	Excellent (85+)
Analyzed Passive	72.86	20.34	Good (72-85)
Summary	49.02	25.90	Poor (39-51)
Correlations	59.38	23.12	OK (51-72)

A one-way between subjects ANOVA analysis showed significant differences between usability for different graph types ($F(4,23)=10.92$, $p<.001$). Post-hoc comparisons between the usability of different graphs showed that the survey response graph was rated significantly more usable that all four other graphs (vs. Data Quality: $p<.001$; vs. Analyzed Passive: $p<.001$; vs Summary: $p<.001$; vs Correlations: $p<.001$). Analyzed Passive graphs were more rated as significantly more usable than the Summary ($p<.001$) or Correlation ($p=.001$) graphs. The Data Quality graph was also rated as more usable than the Summary ($p<.001$) or Correlation ($p=.017$) graphs, and the Correlation graph was only significantly more usable than the Summary graph ($p=.048$).

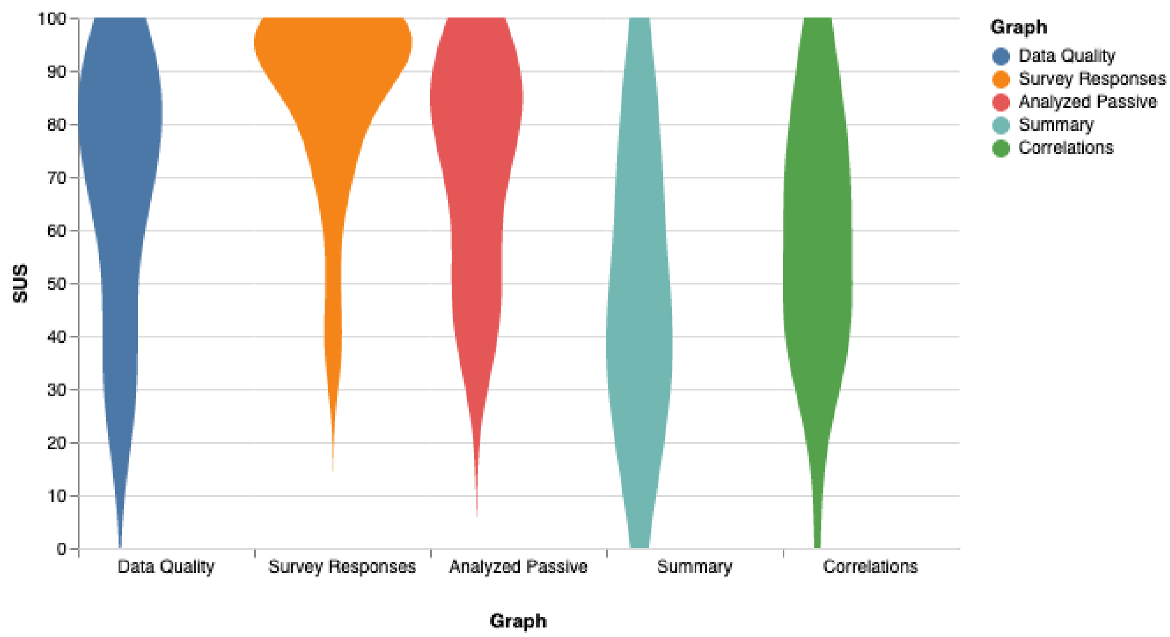

Figure 7. Distribution of SUS Scores, Faceted by Graph Type. This violin plot shows the approximate distributions of different usability scale total scores for each of the graphs shown to participants, faceted by graph type. Wider areas on each distribution correspond to more frequent responses, while narrower areas correspond to fewer or no responses. Higher scores indicate higher usability ratings.

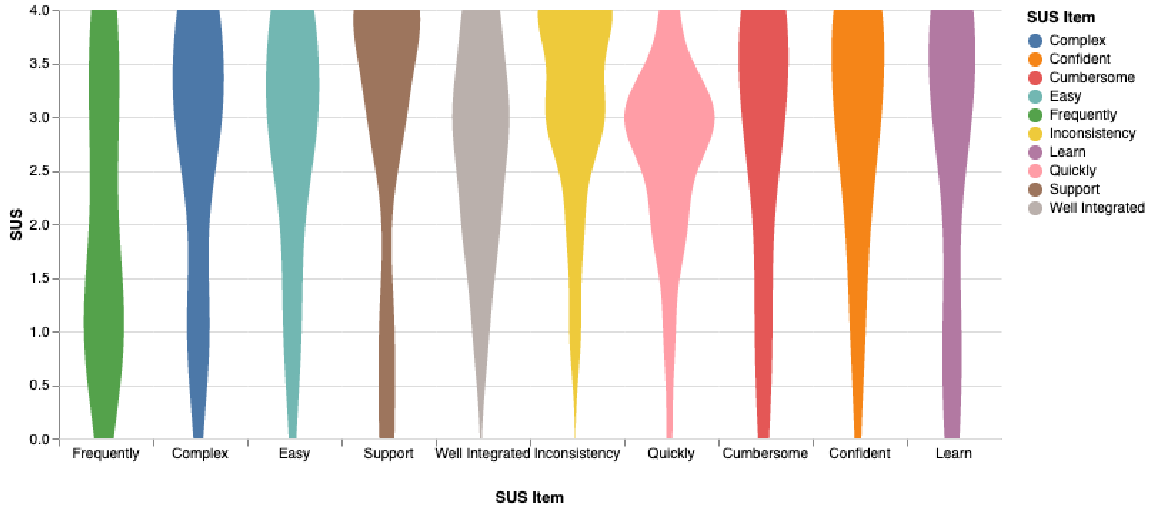

Figure 8. Distribution of SUS Item Scores for Data Quality Graph. This violin plot shows the distribution of scores faceted by each SUS item for the data quality graph. Wider areas on the plots correspond to more frequent responses. Higher scores indicate higher usability.

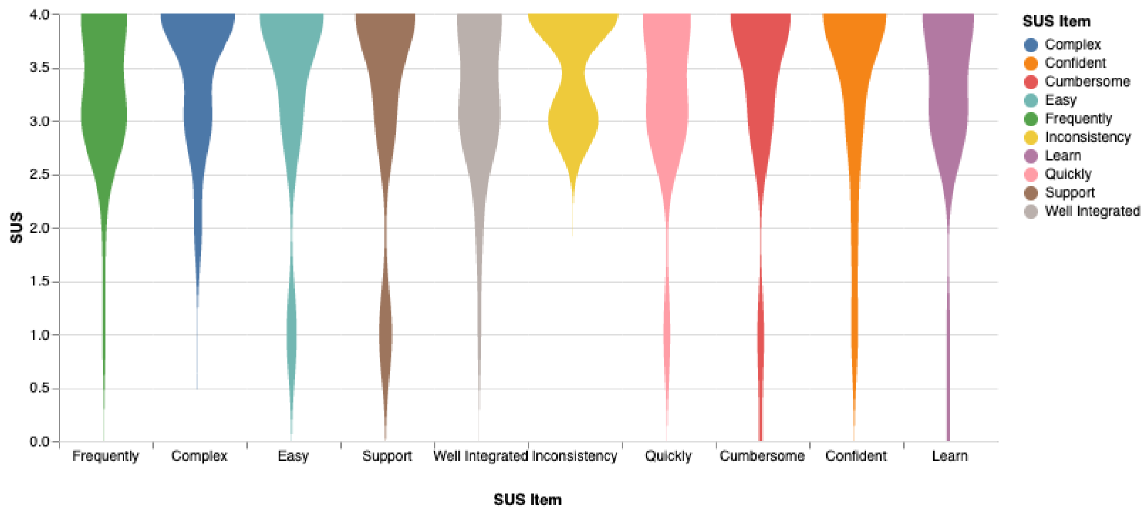

Figure 9. Distribution of SUS Item Scores for Survey Responses Graph. This violin plot shows the distribution of scores faceted by each SUS item for the

survey responses graphs. Wider areas on the plots correspond to more frequent responses. Higher scores indicate higher usability.

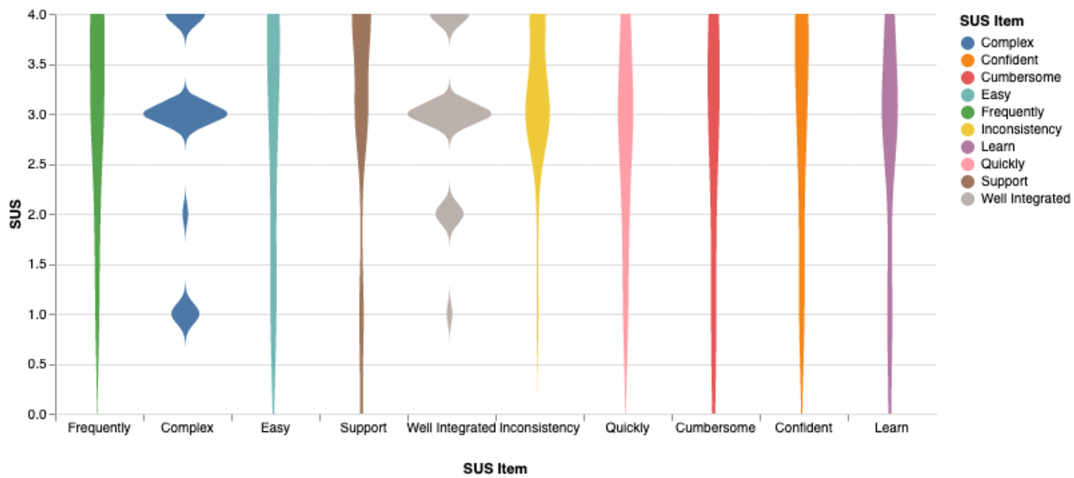

Figure 10. Distribution of SUS Item Scores for Analyzed Passive Graph.

This violin plot shows the distribution of scores faceted by each SUS item for the analyzed passive graph. Wider areas on the plots correspond to more frequent responses. Higher scores indicate higher usability.

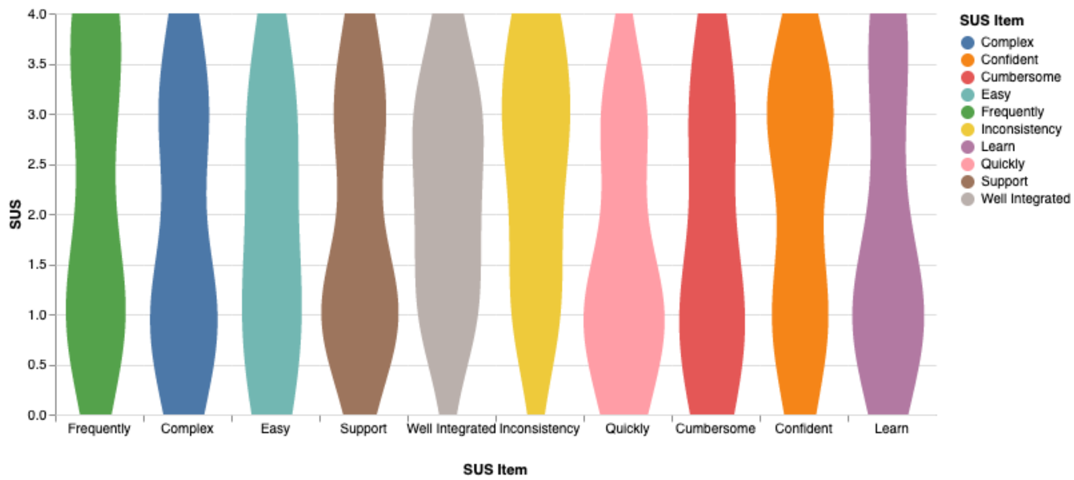

Figure 11. Distribution of SUS Item Scores for Summary Graph. This violin plot shows the distribution of scores faceted by each SUS item for the summary graph. Wider areas on the plots correspond to more frequent responses. Higher scores indicate higher usability.

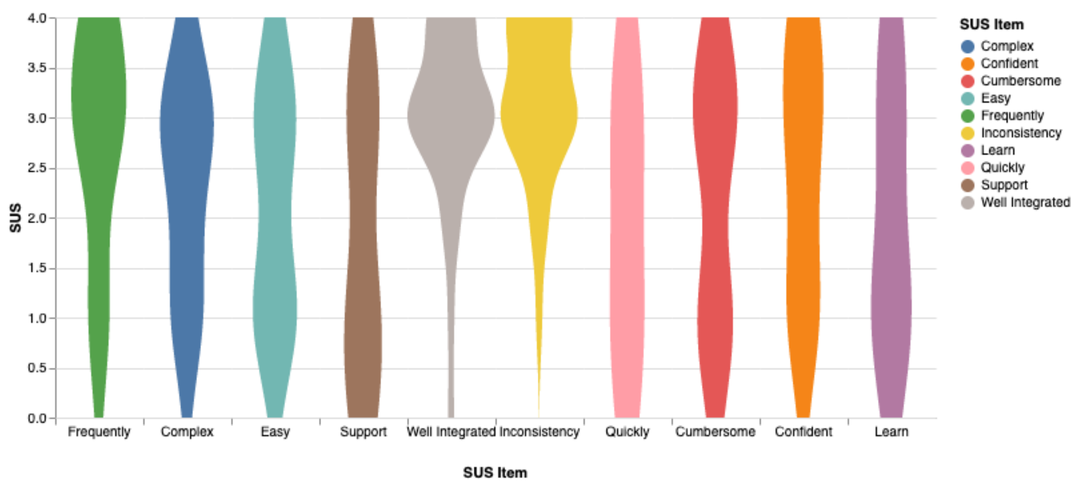

Figure 12. Distribution of SUS Item Score for Correlation Graph. This violin plot shows the distribution of scores faceted by each SUS item for the correlation graph. Wider areas on the plots correspond to more frequent responses. Higher scores indicate more usability.

Additionally, the results of the preferred device type and frequency data were analyzed. There was no statistical difference between graph types for either preferred device ($F(4,23)=0.53$, $p=0.712$) or desired frequency of viewing ($F(4,23)=1.06$, $p=0.377$).

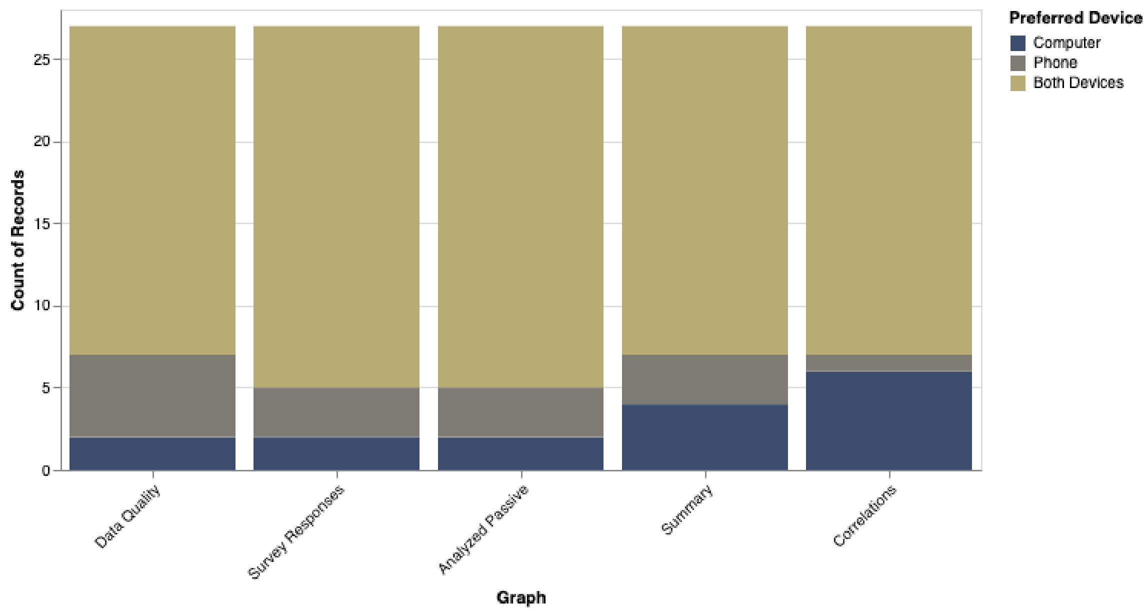

Figure 13. Preferred Device Type for Viewing Visualizations. This graph shows the distribution of different user preferences for how data visualizations would be made available, choosing between a phone or table, computer, or both types of devices.

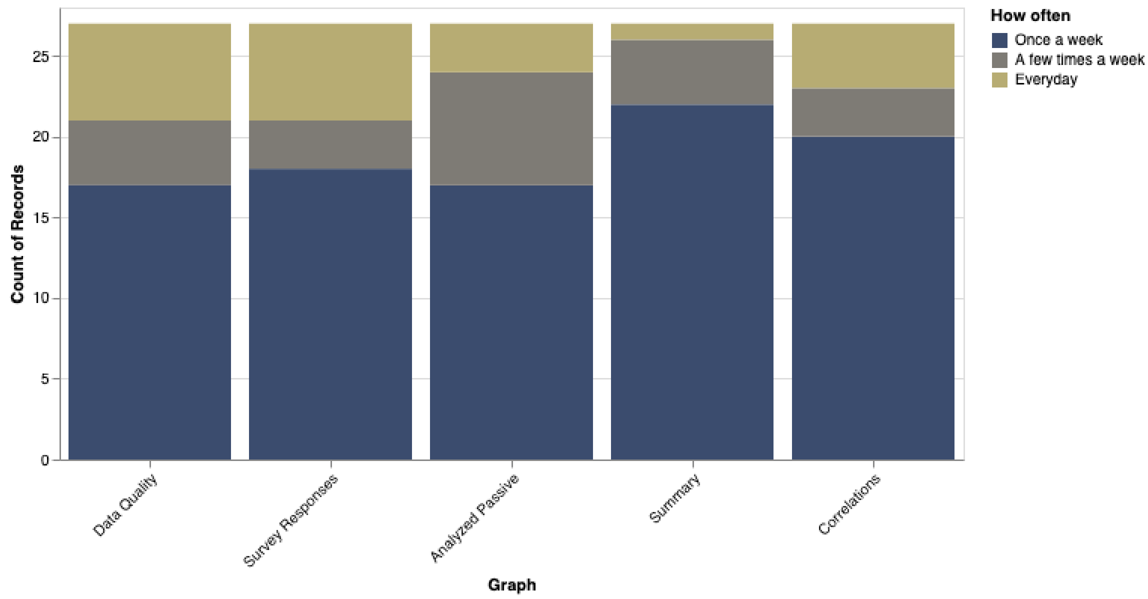

Figure 14. Preferred User Frequency for Viewing Graphs. This graph shows the preferred frequency for how often users wanted to see each of the graphs shown.

Section Three – Digital Working Alliance inventory

Digital Working Alliance Inventory scores, as well as the two added questions, were directly mapped from a scale of “Strongly Disagree”, “Disagree”, “Neither Agree nor Disagree”, “Agree”, or “Strongly Agree” to a 0-4 point scale. No items on the D-WAI are reverse-scored, so higher scores indicate a greater alliance between the user and the graphs.

Table 8. D-WAI and Added Questions Mean Scores by Item. Asterisks indicate the additional items not present in the traditional D-WAI. Higher scores indicate a higher working alliance.

Statement	**Mean (SD)**
I trust these graphs to guide me towards my personal goals	3.14 (0.74)
I believe the graphs would help me to address my problem	3.25 (0.83)
The graphs encourage me to accomplish tasks and make progress	3.11 (0.86)
I agree that using the graphs are important for my goals	3.11 (0.86)
The graphs are easy to use and understand	2.82 (0.92)
The graphs support me to overcome challenges	2.57 (1.05)
These graphs give me a new way to look at my problems*	3.50 (0.68)
These graphs would help me communicate better with my clinician*	3.38 (0.86)

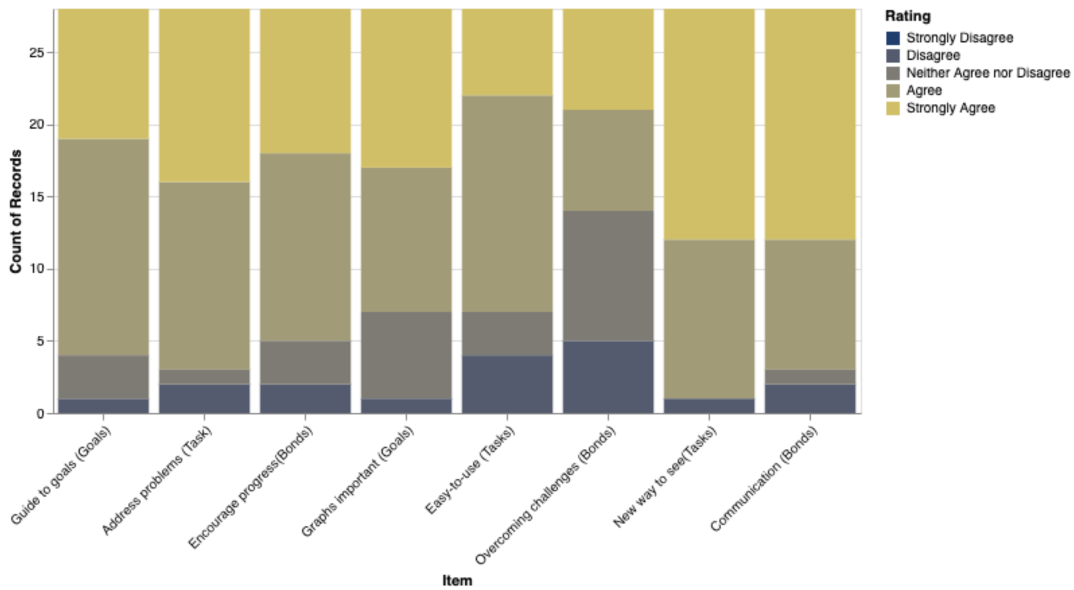

Figure 15. Stacked Bar Chart of DWAI Records. The first six items from left to right represent the traditional D-WAI items, with “New way to see” and “Communication” representing “These graphs give me a new way to look at my problems” and “These graphs would help me communicate better with my clinician” respectively. The words in parentheses represent the three specific Working Alliance Inventory sub-scales – Goal, Task, and Bond. No item was given a rating of “Strongly Disagree” by any participant.

DISCUSSION

Section One - Conclusions

Data visualization remains a promising but largely unexplored means to help patients better understand, engage with, and benefit from digital health data, particularly in clinical settings or when working with healthcare professionals. Our findings indicate that effective data visualizations can change people's willingness to share different kinds of data with their clinicians, help patients better understand their own data, and offer new modalities of identifying and addressing problems, visualizing goals, and communicating with clinicians.

We found data visualization to be a potentially useful method to help people both better understand what data they are sharing, as well as increase their comfort sharing said data. While participants were initially unwilling to share GPS data, ranking it similarly to keylogging data, their perception for GPS changed after viewing the visualizations (Figures 3, 4, and 5), and we observed a statistically significant increase in willingness to share GPS data (Figure 6). This indicates that exposure to the way in which their data will be used and knowledge of how it may factor into their clinical care could predispose users to voluntarily share their data – in other words, patients who understand that GPS data will not be used to track their exact location, but instead to give information about, for example, time spent at home, might feel more comfortable sharing that information with their care team.

Furthermore, we observed the necessary role of education for this purpose; participants' willingness to share keylogging data – which they initially were even less comfortable sharing than GPS data – did not increase after viewing the graphs, which used no measures derived from keylogging data. In fact, average comfort with sharing keylogging data decreased, although not significantly. This indicates that education and training about the uses of data may be necessary to make users feel more comfortable sharing data that they view as sensitive – merely asking the question again is insufficient.

Overall, participants were comfortable sharing most forms of data, which did not largely change after viewing the graphs. After viewing graphs, the combined score across all comfort measures increased from 15.42 out of 20 to 15.89, but this change was not statistically significant. Of particular note is that users were the most comfortable sharing data sourced from surveys – not a single user felt uncomfortable or even neutral about sharing this data before or after looking at the graphs (Figure 6), and all participants rated surveys as at least tied for the form of data they were most comfortable sharing. As discussed in the introduction, this shows that measurement-based care can be well supported by visualizations – participants are already willing to share questionnaire data with care teams and found graphs showing survey data to be highly usable. This means graphs using data from measurement-based care surveys can be easily incorporated into existing care models and electronic medical records, where they become more accessible to clinicians.

In general, users found simpler graphs were more likely to be usable. Participants rated simple graphs that showed, for example, how survey scores on a standard index like the Patient Health Questionnaire changed over time (Figure 4), or how steps taken change day by day (Figure 5), as highly usable, and few if any users rated these graphs as difficult to use (Table 4, Figure 2). However, more complicated graphs like the Summary graph or Correlation charts (Figures 6, Figure 7), that integrated several data metrics were rated as comparatively less usable. This suggests that there is an upper limit to the amount of content that can be shown in a single graph before it becomes confusing

Our results also suggest that interactivity features can make a noticeable difference in usability. Two of the graphs shown, Survey Responses and Correlations contained examples of tooltips whereby a hypothetical user could hover over specific sections of a graph and receive more information; for instance, a record of their responses to particular questions for the survey responses graph, or a description of correlated variables and a brief interpretation of what a correlation value means (Figure 2 and Figure 5). Compared with graphs of similar complexity, Analyzed Passive and Summary, respectively, the graphs with tooltips received higher usability scores (Table 2). In addition, though both the Summary and Correlation graphs were both complex graphs, the Correlation graph, which included an example of an interactive tooltip, was rated as significantly more usable that the summary graph. In particular, participants noted that the features of the Correlation graph were both better integrated and more consistent than the features of the Summary graph (Figure 11 and Figure 12). This suggests that adding interactivity to data visualizations through a system like tooltips can make both simple data more detailed and complex data more understandable; this provides a potential target for additional future research.

Graphs need to clearly demonstrate their use case for participants to take an interest in them – in addition, participants are able to quickly identify graphs they do not see a use for. While the data quality graph, which showed participants how much data they were collecting received similar scores to the more usable graphs in metrics like how quickly participants felt they could learn to use the graphs, it received much lower ratings for how frequently participants would want to use it, and the overall distribution of scores was much wider (Figure 6, Figure 7, Figure 8, Figure 9). Clinicians or researchers might see the value in a graph that shows the quality of a user's data – for example, it is important to ensure that any correlations or other metrics calculated, such as hometime or screen usage, are of an appropriately quality and accurately reflect the true state of a participant's actions. However, if a participant doesn't feel that data quality information is useful, they may not want to use or monitor it – and may miss information like low data quality that they could otherwise correct. This would likely extend to other forms of data sharing and visualizations – if participants do not perceive value in certain forms of graphs, they are unlikely to want to use them, highlighting the need for clinicians to effectively communicate the purpose of different data visualizations to ensure compliance.

Interestingly, while participants saw differences in the usability of graphs, there was little variation in how frequently they wanted to see these graphs – an ANOVA test of the desired frequency between graph types revealed no significant differences between how often participants wanted to use these graphs in a clinical setting (Figure 14). This suggests that the average participant is not currently interested in constant symptom monitoring; most want to see even graphs they find usable or desirable no more than once a week, and only a small subset want to use the graphs every single day. Thus, if a clinician wants a patient to play an active, everyday role in monitoring their mental status, that will need to be an active topic of conversation in early meetings.

Most participants preferred maximal flexibility in terms of how they could view graphs and visualizations, with a large majority of participants preferring to be able to view graphs on both their phone and computer. These results make sense as national data suggests that both phones and computers are being used by patients at nearly equal rates to access telehealth (Roberts & Mehrotra, 2020) However, they also present a challenge as a visualization optimized for a smartphone may be different from ones optimized for the larger display afforded by a computer. The visualizations we shared in the study are flexible and can be used across both types of devices given the goal to support patients across as many devices as possible.

Participants were able to identify graphs they valued both for use on their own, as well as graphs that they wanted to use in concert with a clinician. After adjusting for differences in scoring, the graph system used for this study scored similarly, slightly better in fact, on the D-WAI when compared to the average D-WAI ratings found by another study of people's most commonly used meditation apps, 33.87 vs. 30.58 (Goldberg et al., 2021). This suggests that the participants of this study felt they would be able to successfully work with these apps in a clinical setting. While simple graphs were rated as more independently usable than complex ones, most participants saw the value of all the graphs-- of the 28 participants surveyed, 27 agreed that using the graphs would give them "a new way to look at their problems," and 25 agreed that the graphs would help them "communicate better with [their] clinician" (Figure 15). This indicates that users still saw the value in complex summary or correlation graphs they rated as less usable and understood the intermediary role that physicians, other clinicians, and digital navigators could play to help patients get the greatest value out of their data in a clinical setting. While beyond the scope of this thesis, this point raises the issue of ensuring that clinicians have the appropriate training and feel comfortable discussing digital data in care.

Section Two - Limitations

Our study has some limitations. First, a relatively small sample size of 28, which reflects the nature of this as both a pilot study and one conducted under a more limited time frame. Second, while the fact that participants in this study were all sampled from a larger study where an entry criteria was scoring moderately or above on Perceived Stress Scale is an advantage -- as the studied issues are particularly relevant to a clinical population -- the fact that the larger study from which these participants were drawn examined the uses of digital technology could mean that participants who completed that study were more likely to be willing to share data, may find technology more usable than the average person, and may be more likely to want to use technology in a clinical setting than may be true for an average patient. Third, we only collected data on participants' opinions regarding sharing data in a clinical setting, and not actual usage and behaviors around data sharing and clinical interactions, or whether participants would share their data the same way in a research setting as well.

Section Three – Future Directions

This work also suggests several potential next steps to continue investigating how data visualizations can successfully be integrated into clinical care.

First, we identified the important role clinicians could play in helping patients to understand data when it must be presented in a more complicated way, like a summary or correlation graph. A next step, then, is to meet with clinicians and conduct similar usability analysis to learn what they value in data visualizations and if they see value in having graphs available to them and their patients. This will allow for the creation of focused training to help ensure clinicians have the knowledge, skills, and confidence to incorporate digital data into care.

Second, our results for the effect of interactivity on visualization usability was encouraging; usability scores for graphs with tooltips were around 10 points higher compared to graphs of similar complexity that did not utilize tooltips. This suggests directly comparing the usability of the same graph with and without tooltips could further tease out the direct effects of interactivity, which might help make more complicated graphs like the Summary and Correlation charts usable by those without clinical backgrounds.

Third, while examining a population who scored at least moderately on a measure of symptoms of stress increases the ability of this study to inform decisions in a clinical setting, it would be good to examine how data visualizations are perceived both by healthy controls as well as individuals with more severe forms of mental illness to get a more complete picture of how a range of individuals could use these visualizations.

Investigating these issues will provide valuable information and help create a more complete system of visualizations that support both patient and clinician.

JOURNAL ABBREVIATIONS

BMJ	British Medical Journal
BJPsych	The British Journal of Psychiatry
JAMA	Journal of the American Medical Association
JMIR	Journal of Medical Internet Research
J of ACH	Journal of American College Health
JUS	Journal of Usability Studies

REFERENCES

Adanijo, A., McWilliams, C., Wykes, T., & Jilka, S. (2021). Investigating Mental Health Service User Opinions on Clinical Data Sharing: Qualitative Focus Group Study. *JMIR Mental Health*, *8*(9), e30596. https://doi.org/10.2196/30596

Bangor, A., Kortum, P. T., & Miller, J. T. (2008). An Empirical Evaluation of the System Usability Scale. *International Journal of Human–Computer Interaction*, *24*(6), 574–594. https://doi.org/10.1080/10447310802205776

Bangor, A., Kortum, P. T., & Miller, J. T. (2009). Determining what individual SUS scores mean: Adding an adjective rating scale. *JUS - The Journal of Usability Studies*. https://uxpajournal.org/determining-what-individual-sus-scores-mean-adding-an-adjective-rating-scale/

Bauer, A. M., Iles-Shih, M., Ghomi, R. H., Rue, T., Grover, T., Kincler, N., Miller, M., & Katon, W. J. (2018). Acceptability of mHealth augmentation of Collaborative Care: A mixed methods pilot study. *General Hospital Psychiatry*, *51*, 22–29. https://doi.org/10.1016/j.genhosppsych.2017.11.010

Baumel, A., Muench, F., Edan, S., & Kane, J. M. (2019). Objective User Engagement With Mental Health Apps: Systematic Search and Panel-Based Usage Analysis. *Journal of Medical Internet Research*, *21*(9), e14567. https://doi.org/10.2196/14567

Bilden, R., & Torous, J. (2022). Global Collaboration Around Digital Mental Health: The LAMP Consortium. *Journal of Technology in Behavioral Science*, 1–7. https://doi.org/10.1007/s41347-022-00240-y

Dowrick, C., Leydon, G. M., McBride, A., Howe, A., Burgess, H., Clarke, P., Maisey, S., & Kendrick, T. (2009). Patients' and doctors' views on depression severity questionnaires incentivised in UK quality and outcomes framework: Qualitative study. *BMJ (Clinical Research Ed.)*, *338*, b663. https://doi.org/10.1136/bmj.b663

Fortney, J. C., Unützer, J., Wrenn, G., Pyne, J. M., Smith, G. R., Schoenbaum, M., & Harbin, H. T. (2017). A Tipping Point for Measurement-Based Care. *Psychiatric Services*, *68*(2), 179–188. https://doi.org/10.1176/appi.ps.201500439

Gansner, M., Nisenson, M., Carson, N., & Torous, J. (2020). A pilot study using ecological momentary assessment via smartphone application to identify adolescent problematic internet use. *Psychiatry Research*, *293*, 113428. https://doi.org/10.1016/j.psychres.2020.113428

Gansner, M., Nisenson, M., Lin, V., Carson, N., & Torous, J. (2022). Piloting Smartphone Digital Phenotyping to Understand Problematic Internet Use in an Adolescent and Young Adult Sample. *Child Psychiatry and Human Development*. https://doi.org/10.1007/s10578-022-01313-y

Glomann, L., Hager, V., Lukas, C. A., & Berking, M. (2019). Patient-Centered Design of an e-Mental Health App. In T. Z. Ahram (Ed.), *Advances in*

Artificial Intelligence, Software and Systems Engineering (pp. 264–271). Springer International Publishing. https://doi.org/10.1007/978-3-319-94229-2_25

Goldberg, S. B., Baldwin, S. A., Riordan, K. M., Torous, J., Dahl, C. J., Davidson, R. J., & Hirshberg, M. J. (2021). Alliance With an Unguided Smartphone App: Validation of the Digital Working Alliance Inventory. *Assessment*, 10731911211015310. https://doi.org/10.1177/10731911211015310

Henson, P., Pearson, J. F., Keshavan, M., & Torous, J. (2020). Impact of dynamic greenspace exposure on symptomatology in individuals with schizophrenia. *PloS One*, *15*(9), e0238498. https://doi.org/10.1371/journal.pone.0238498

Henson, P., Wisniewski, H., Hollis, C., Keshavan, M., & Torous, J. (2019). Digital mental health apps and the therapeutic alliance: Initial review. *BJPsych Open*, *5*(1), e15. https://doi.org/10.1192/bjo.2018.86

Iliescu, R., Kumaravel, A., Smurawska, L., Torous, J., & Keshavan, M. (2021). Smartphone ownership and use of mental health applications by psychiatric inpatients. *Psychiatry Research*, *299*, 113806. https://doi.org/10.1016/j.psychres.2021.113806

Jaworski, B. K., Taylor, K., Ramsey, K. M., Heinz, A., Steinmetz, S., Pagano, I., Moraja, G., & Owen, J. E. (2021). Exploring Usage of COVID Coach, a Public Mental Health App Designed for the COVID-19 Pandemic:

Evaluation of Analytics Data. *Journal of Medical Internet Research*, *23*(3), e26559. https://doi.org/10.2196/26559

Kroenke, K., & Spitzer, R. L. (2002). The PHQ-9: A new depression diagnostic and severity measure. *Psychiatric Annals*, *32*(9), 509–515. https://doi.org/10.3928/0048-5713-20020901-06

Lagan, S., D'Mello, R., Vaidyam, A., Bilden, R., & Torous, J. (2021). Assessing mental health apps marketplaces with objective metrics from 29,190 data points from 278 apps. *Acta Psychiatrica Scandinavica*, *144*(2), 201–210. https://doi.org/10.1111/acps.13306

Lewis, C. C., Boyd, M., Puspitasari, A., Navarro, E., Howard, J., Kassab, H., Hoffman, M., Scott, K., Lyon, A., Douglas, S., Simon, G., & Kroenke, K. (2018). Implementing Measurement-Based Care in Behavioral Health: A Review. *JAMA Psychiatry*, 10.1001/jamapsychiatry.2018.3329. https://doi.org/10.1001/jamapsychiatry.2018.3329

Luo, H., Lee, P.-A., Clay, I., Jaggi, M., & De Luca, V. (2020). Assessment of Fatigue Using Wearable Sensors: A Pilot Study. *Digital Biomarkers*, *4*(1), 59–72. https://doi.org/10.1159/000512166

Melcher, J., Lavoie, J., Hays, R., D'Mello, R., Rauseo-Ricupero, N., Camacho, E., Rodriguez-Villa, E., Wisniewski, H., Lagan, S., Vaidyam, A., & Torous, J. (2021). Digital phenotyping of student mental health during COVID-19: An observational study of 100 college students. *Journal of American*

College Health: J of ACH, 1–13. https://doi.org/10.1080/07448481.2021.1905650

Mote, J., & Fulford, D. (2019). Increasing access to evidence-based care in serious mental illness: Embracing mobile technology while minding the digital divide. *The Clinical Psychologist*, *72*, 5–12.

NW, 1615 L. St, Washington, S. 800, & Inquiries, D. 20036 U.-419-4300 | M.-857-8562 | F.-419-4372 | M. (2021, April 7). Demographics of Mobile Device Ownership and Adoption in the United States. *Pew Research Center: Internet, Science & Tech*. https://www.pewresearch.org/internet/fact-sheet/mobile/

Parker, L., Halter, V., Karliychuk, T., & Grundy, Q. (2019). How private is your mental health app data? An empirical study of mental health app privacy policies and practices. *International Journal of Law and Psychiatry*, *64*, 198–204. https://doi.org/10.1016/j.ijlp.2019.04.002

Polhemus, A., Novák, J., Majid, S., Simblett, S., Bruce, S., Burke, P., Dockendorf, M., & Temesi, G. (2020). *Data visualization in chronic neurological and mental health condition self-management: A systematic review of user perspectives (Preprint)*. https://doi.org/10.2196/preprints.25249

Possemato, K., Kuhn, E., Johnson, E., Hoffman, J. E., Owen, J. E., Kanuri, N., De Stefano, L., & Brooks, E. (2016). Using PTSD Coach in primary care with and without clinician support: A pilot randomized controlled trial.

General Hospital Psychiatry, *38*, 94–98. https://doi.org/10.1016/j.genhosppsych.2015.09.005

Quaedackers, L., Wit, J. D., Pillen, S., Gilst, M. V., Batalas, N., Lammers, G. J., Markopoulos, P., & Overeem, S. (2020). A Mobile App for Longterm Monitoring of Narcolepsy Symptoms: Design, Development, and Evaluation. *JMIR MHealth and UHealth*, *8*(1), e14939. https://doi.org/10.2196/14939

Rauseo-Ricupero, N., Henson, P., Agate-Mays, M., & Torous, J. (2021). Case studies from the digital clinic: Integrating digital phenotyping and clinical practice into today's world. *International Review of Psychiatry*, *33*(4), 394–403. https://doi.org/10.1080/09540261.2020.1859465

Reay, R. E., Looi, J. C., & Keightley, P. (2020). Telehealth mental health services during COVID-19: Summary of evidence and clinical practice. *Australasian Psychiatry*, *28*(5), 514–516. https://doi.org/10.1177/1039856220943032

Roberts, E. T., & Mehrotra, A. (2020). Assessment of Disparities in Digital Access Among Medicare Beneficiaries and Implications for Telemedicine. *JAMA Internal Medicine*, *180*(10), 1386–1389. https://doi.org/10.1001/jamainternmed.2020.2666

Simblett, S., Greer, B., Matcham, F., Curtis, H., Polhemus, A., Ferrão, J., Gamble, P., & Wykes, T. (2018). Barriers to and Facilitators of Engagement With Remote Measurement Technology for Managing

Health: Systematic Review and Content Analysis of Findings. *Journal of Medical Internet Research*, *20*(7), e10480. https://doi.org/10.2196/10480

Singh, S., Strong, R. W., Jung, L., Li, F. H., Grinspoon, L., Scheuer, L. S., Passell, E. J., Martini, P., Chaytor, N., Soble, J. R., & Germine, L. (2021). The TestMyBrain Digital Neuropsychology Toolkit: Development and Psychometric Characteristics. *Journal of Clinical and Experimental Neuropsychology*, *43*(8), 786–795. https://doi.org/10.1080/13803395.2021.2002269

Spitzer, R. L., Kroenke, K., Williams, J. B. W., & Löwe, B. (2006). A Brief Measure for Assessing Generalized Anxiety Disorder: The GAD-7. *Archives of Internal Medicine*, *166*(10), 1092–1097. https://doi.org/10.1001/archinte.166.10.1092

Steinfeld, B., Franklin, A., Mercer, B., Fraynt, R., & Simon, G. (2016). Progress Monitoring in an Integrated Health Care System: Tracking Behavioral Health Vital Signs. *Administration and Policy in Mental Health and Mental Health Services Research*, *43*(3), 369–378. https://doi.org/10.1007/s10488-015-0648-7

Substance Abuse and Mental Health Services Administration. (2021). *Key Substance Use and Mental Health Indicators in the United States: Results from the 2020 National Survey on Drug Use and Health* (p. 156) [HHS Publication No. PEP21-07-01-003, NSDUH Series H-56]. Center for

Behavioral Health Statistics and Quality, Substance Abuse and Mental Health Services Administration. https://www.samhsa.gov/data/

Torous, J., Friedman, R., & Keshavan, M. (2014). Smartphone ownership and interest in mobile applications to monitor symptoms of mental health conditions. *JMIR MHealth and UHealth*, *2*(1), e2. https://doi.org/10.2196/mhealth.2994

Vaidyam, A., Halamka, J., & Torous, J. (2022). Enabling Research and Clinical Use of Patient-Generated Health Data (the mindLAMP Platform): Digital Phenotyping Study. *JMIR MHealth and UHealth*, *10*(1), e30557. https://doi.org/10.2196/30557

Wisniewski, H., Henson, P., & Torous, J. (2019). Using a Smartphone App to Identify Clinically Relevant Behavior Trends via Symptom Report, Cognition Scores, and Exercise Levels: A Case Series. *Frontiers in Psychiatry*, *10*, 652. https://doi.org/10.3389/fpsyt.2019.00652

VITA

[redacted]

[redacted]

[redacted]

[redacted]