Nottingham Trent University

Evaluation of the Improving Access to Psychological Therapies (IAPT) Chatbot with Trent Psychological Therapy Services

Dr Maria-Karanika-Murray, Dr Gabriella Mutale, Professor Di Bailey and Professor Steven Brown

January 2022

Table of Contents

Executive Summary

Chatbots are digital tools that can use machine learning and artificial intelligence methods to mimic humanlike responses and behaviours in a conversation. Over the last 5 years chatbots have become increasingly popular in mental health and have been used to provide therapy and other forms of mental health support. Research suggests that the use of chatbots in mental health provision may be positive however, there is limited evidence making it difficult to determine their usefulness and effectiveness.

The Improving Access to Psychological Therapies (IAPT) programme began in 2008 and has transformed the treatment of mental health conditions in England by allowing easier access to therapy for millions with common mental health conditions. The NHS Mental Health Implementation Plan states that IAPT services, in line with the NHS's vision for digital healthcare, should offer a range of self-management apps, digital consultations and digitally-enabled models of therapy to support treatment by 2023/24.

Trent Psychological Therapy Service (PTS) is an NHS approved provider of the IAPT service in Nottinghamshire and Derbyshire. In 2021 Trent PTS partnered with Virtual Health Labs to design and launch a chatbot to be used with people who use Trent PTS. The chatbot was designed as an automated conversational agent to help increase the reach and retention of those on the IAPT treatment pathway. In total four different chatbots were implement by Trent PTS for use with their clients*:

Chatbot	Description
DNA chatbot (launched March 2021)	Offered to clients who do not attend (DNA) their first assessment. Aims to establish if there is a problem with engagement and encourages clients to re-book another appointment
Step up chatbot (launched March 2021)	Offered to clients who are on the waiting list for Step 3 therapy. The chatbot gives suggestions of things the client could do whilst they wait for first therapy session
Solution chatbot (launched October 2021)	Offered to clients before their first appointment. Intended to get clients to think about a current problem they have and takes them through brief solution focused coaching enabling them to identify small steps they can take to improve their situation
Prepared Client chatbot (launched October 2021)	Offered to clients before their first appointment. It aims to prepare the client for their therapy sessions asking them about values, concerns and hopes for therapy. There is the option to forward the conversation to the therapist.

Nottingham Trent University were commissioned by the East Midlands Academic Science Health Network to undertake an evaluation of the chatbots being used by Trent PTS. The aims of the evaluation were to understand:

- What are the enables and barriers to implementing a chatbot?
- How does a chatbot impact on the IAPT service from staff perspective?
- What is the relationship that clients develop with a chatbot?
- Is client satisfaction good when using a chatbot?
- What effect does a chatbot have on retention, attendance and dropout rates?

The evaluation used a multi-level and multi method framework (e.g., Bailey et al., 2017, 2020; Bailey & Kerlin, 2015; Bailey & Mutale, 2020; Mutale et al., 2020; Ward & Bailey, 2015) which included collecting both quantitative and qualitative data across several levels (context, inputs, and outcomes). Data sources included were:

- Demographic data from clients using chatbots
- Data on the use and uptake of chatbots
- Data from focus group/interview with staff members
- Client feedback data collected from chatbots and a client survey
- Data on clients' attendance at appointments

Key Findings

Level of Evaluation	Key Findings
Context **What are the enablers and barriers to implementing the chatbot?**	It was found that when implementing a chatbot the client group which is to be targeted needs to be considered and included in the design of the chatbot. The evaluation found that clients who do not attend (DNA) their assessment were not receptive to engaging with the DNA chatbot suggesting the DNA chatbot may not be an effective way to target this client group. Chatbots that were launched further into the evaluation (Solution Chatbot, Prepared Client Chatbot) had greater success. These chatbots focused on helping clients with problems and preparing them for therapy, suggesting clients are more likely to engage with a chatbot when it is more explicitly linked to the therapy process.
Inputs **How does the chatbot impact on the IAPT service from staff perspective?**	Staff responded that clients who do not attend appointments were not the ideal group to target with a chatbot. Staff participating in the evaluation suggested that the chatbot should be offered to all clients at the point of referral so that they could then choose if they wanted to use it, with the recognition that some clients will not want to engage with a chatbot. Staff felt the option of using a chatbot could be beneficial if used throughout the IAPT journey for clients.
Outcomes	Feedback collected from the chatbots suggested that clients mainly develop a positive relationship with the chatbot. Clients

What is the relationship that service users develop with the chatbot? **Is client satisfaction good when using the chatbot?** **What effect does the chatbot have on retention, attendance and drop out rates?**	using the Solution chatbot reported that the chatbot was able to help them think more clearly about their problems, reduced anxiety and helped them to feel calmer and provided them with ideas of how to help themselves. Clients also felt being able to talk about their problems with a chatbot was positive. In relation to preparing for their first appointment clients felt that the prepared client chatbot was successful in enabling them to do this and thought the process was useful. Some clients enjoyed the anonymity talking to the chatbot provided. Clients using the Solution chatbot had some concerns about talking to a chatbot rather than a therapist and feeling left alone. This could be counteracted if in future the chatbot is used alongside therapy integrating it more into the therapy process. It is possible that clients who continue to interact with the chatbot over a longer period of time may feel the process becomes easier. Client satisfaction was good for the majority (over half) of clients with clients leaving positive feedback. The findings suggests that some individuals do not want to use a chatbot and feedback suggests this is because they find it hard to talk to a robot finding the responses too generic and not tailored to their own needs. The DNA chatbot did not reduce further DNAs suggesting that this chatbot is ineffective in overcoming DNAs. This further highlights the problems with using a chatbot with this client group. Further research is needed to understand clients' reasons for not attending appointments. The Solution chatbot was effective at reducing cancellations and DNAs when compared to the Prepared Client chatbot and controls. The Solution chatbot saw a 19.44% decrease in DNAs and a 38.89% decrease in cancellations compared to controls. Clients using the Prepared Client chatbot had a 36.11% increase in DNAs and a 16.67% increase in cancellations compared to controls. More research is needed to understand why the Solution chatbot is able to encourage attendance at appointments and to examine how the Prepared Client chatbot can be developed to achieve this. The evaluation highlights that a chatbot is better used to help clients before their first appointment where the chatbot content directly relates to the client's presenting problem or the therapy process. This would suggest that a chatbot could continue to be used with clients alongside the therapy process in a positive way that maximises engagement. Further evaluation is needed to assess how successful this could be and to establish the best way to integrate the chatbot into therapy.

*Through this report we refer to the people who use Trent PTS as *clients*. This is in line with the language Trent PTS use to refer to the people who use their service. Other services may use the terms patients or service users and within the literature these terms are often used.

1. Introduction

1.1 What are Chatbots?

Chatbots are digital tools existing either as hardware or software that can use machine learning and artificial intelligence methods to mimic humanlike behaviours which allows the users to participate in conversation with the chatbot (Vaidyam et al., 2019). Chatbots are accessible to anyone with a smartphone or computer and internet access. Conversational agents and chatbots, such as Apple's Siri or Amazon's Alexa, have become increasingly popular in the digital marketplace There are 2 types of chatbots: 1) rule-based chatbots which use pre-defined rules or decision trees to create a response and 2) intelligent chatbots which use artificial intelligence to produce responses (Abd-Alrazaq et al., 2021).

Although Chatbots have been used in various industries such as customer services, retail, travel, and entertainment it is only over the last five years they have become popular in mental health provision (Abd-Alrazaq et al., 2021). It has been shown that there is potential for chatbots to provide mental health advice and support. For example, research (Miner et al., 2016) examining the response of four popular smartphone conversational agents (Apple's Siri, Google Now, Samsung Voice and Microsoft's Cortana) response to questions related to mental health found that when asked about committing suicide three of the agents were able to recognise the concern and respond in a respectful way, e.g., by directing the user towards a helpline and showing concern. When asked about depression all of the conversational agents were able to recognise this and respond in a respectful way (e.g., 'I'm very sorry. Maybe it would help to talk to someone'). This shows the ability of chatbots to be able to recognise and respond to mental health concerns, thus highlighting their potential for providing mental health support.

1.2 How Effective are Chatbots When Used in Mental Health?

There has been a lack of evidence on the use of chatbots in mental health. A systematic review of the literature conducted by Vaidyam et al. (2019) identified some of the benefits of the use of chatbots in mental health. These were high patient satisfaction levels, low risk of harm and the provision of psychoeducation. The review showed that chatbots give the potential for clients to provide self-care. Only one randomised controlled trial (RCT) (Fitzpatrick et al., 2017) examined the effect of chatbots this found that the use of a chatbot was able to significantly reduce depressive symptoms in clients.

A further potential benefit of chatbots that has been identified is the anonymity they provide. This means that clients are more willing to disclose sensitive information compared to when receiving therapy from a real-life therapist (Lucas et al., 2014). This was found to be because clients felt that no one was observing or judging them, and consequently that they could talk without fear of negative evaluation. Chatbots could therefore be helpful in providing therapy to clients who feel uncomfortable talking about their feelings with a therapist.

A recent systematic review of the literature (Abd-Alrazaq et al., 2021) found that when providing mental health support chatbots can deliver some elements of support that health care services cannot provide over a sustained period. For example, real time feedback, weekly summaries of progress and continuous data collection. The review found that patient satisfaction with chatbots is generally good, however common issues reported are the conversational limitations of chatbots (Abd-Alrazaq et al., 2021). This may be because chatbots in mental health are still in an early development stage. The review found that the majority of chatbots in mental health are rule based using decision trees to generate their responses. This is lagging behind chatbots used in other areas, for example customer services, where artificial intelligence chatbots are more commonly used.

There have been concerns raised as to whether a chatbot's lack of ability to re-create human interactions could lead to harm when being used in mental health settings. However, there is a lack of research which has evaluated patient safety and this needs to be more specifically tracked (Kretzschmar et al., 2019; Laranjo et al., 2018)

Other research (Bell et al., 2019) has suggested that therapy provided by chatbots does not result in an improved patient reported experience. Bell et al found that participants who had received therapy from a chatbot had lower levels of enjoyment and perceived the session to be less useful than the control group who had therapy from an actual therapist.

Although there has been an increasing amount of research regarding the use of chatbots in mental health the research in this area is still lacking and there is a scarcity of robust evidence that demonstrates the effectiveness of chatbots in mental health. There is some evidence (Abd-Alrazaq et al., 2021; Vaidyam et al., 2019) to indicate that the use of chatbots for mental health provision may result in positive outcomes and user experience but this is limited making it difficult to determine the usefulness and effectiveness of chatbots.

1.3 The Use of Chatbots in the Improving Access to Psychological Therapy Service

The Improving Access to Psychological Therapy (IAPT) programme began in 2008 and has transformed the treatment of mental health conditions such as depression and anxiety (NHS, 2021). The NHS Mental Health Implementation Plan (NHS, 2019) states that by 2023/24 the number of people with mental health conditions that can access psychological therapy through IAPT will be increased to 1.9 million. The plan states that the IAPT service in line with the NHS's vision for digital healthcare should offer a range of self-management apps, digital consultations and digitally-enabled models of therapy to support treatment by 2023/24.

Trent Psychological Therapy Service (PTS) is an NHS approved provider of the IAPT service in Nottinghamshire and Derbyshire. In 2021 Trent PTS launched an IAPT Augmentbot (chatbot) to be used with their clients. The chatbot is an automated conversational agent and has been designed to help increase the reach and retention of clients through the IAPT treatment pathway. The chatbot was designed to influence:

- Attending appointments
- Engaging in therapy
- Completing therapy

1.4 Evaluation Aims

The evaluation was commissioned by the East Midlands Academic Science Health Network to explore how the chatbot was being used by Trent PTS. The aims of the evaluation were to understand:

- What are the enables and barriers to implementing the chatbot?
- How does the chatbot impact on the IAPT service from staff perspective?
- What is the relationship that clients develop with the chatbot?
- Is client satisfaction good when using the chatbot?
- What effect does the chatbot have on retention, attendance and drop out rates?

2. Methodology

2.1 Evaluation Design

The evaluation design was guided by Pawson and Tilley's (2004) realistic evaluation methodology that takes ''heed of the different layers of social reality which make up and surround programmes of change'' (p4). To create our multi-level evaluation framework, we have taken levels from previously developed frameworks by Warr, Bird and Rackham (1970) and Kirkpatrick (1994) and combined them to create a new framework (see Table 1). We have developed and refined our evaluation framework through previous evaluation research (e.g. Bailey et al., 2017, 2020; Bailey & Kerlin, 2015; Bailey & Mutale, 2020; Biron & Karanika-Murray, 2014; Karanika-Murray et al.,2016; Mutale et al., 2020; Ward & Bailey, 2015;). Through this we have demonstrated that our evaluation design and framework can be used successfully to measure the impact of different interventions in a variety of mental health and social care services in a robust way.

The levels of the evaluation framework, research questions and data to be collected are shown in Table 1.

Table 1 Evaluation framework

Level of evaluation	Research questions to be answered	Data sources
Context: In what context is the chatbot used?	What are the enablers and barriers to implementing the chatbot?	• Demographic data on clients using chatbots

		• Uptake of IAPT from clients using chatbots
Inputs: What is the service able to deliver through using the chatbot?	How does the chatbot impact on the IAPT service from staff perspective?	• Focus group/interview with staff
Outcomes: What are the outcomes for Trent PTS and its clients?	Does the chatbot help improve retention, attendance and drop out rates? Is client satisfaction good when using the chatbot? What is the relationship that service users develop with the chatbot?	• Data on clients' attendance at appointments and DNAs • Client survey • Client feedback collected from chatbot

2.2 Chatbot Design and Implementation

Originally two chatbots were designed by Virtual Health Labs in conjunction with Trent PTS with an overall aim of reducing the number of clients who do not attend (DNA) their appointment or drop out while on the waiting list for therapy. These chatbots were:

DNA Chatbot:
Is offered to people who DNA their initial assessment.
It aims to ascertain if there is a problem and to establish if the client would like another appointment.
It is designed to stop clients going on to DNA further appointments.

Things the DNA chatbot might say:

- Reassures the client that missing appointments can happen
- Offers the client a chance to re-book
- If the client is not interested in re-booking it will gently enquire why.

Step up Chatbot:
Is offered to people who are on the waiting list for Step 3 therapy.
It is designed to reduce drop out rates at the stage and gives ideas of what clients could do while they wait for their appointment.

Things the Step Up chatbot might say:

- Gives reassurance to the client
- Offers the client helpful things to do while waiting (e.g. physical exercise, eating well, reduce social media)
- Help if the client is in crisis (e.g. signposted to NHS 111, Samaritans or 999)

The DNA and Step Up chatbots were offered at random to half of the clients who were eligible. After low uptake of the original chatbots two new chatbots were introduced, targeting a different sample. The chatbots still aimed to reduce DNAs and dropouts specifically through targeting different groups. These chatbots were a Solution chatbot and a Prepared Client chatbot. Clients who were waiting for their first appointment were offered at random either a) the Prepared Client chatbot b) the Solution chatbot or c) no chatbot – treatment as usual (control group).

Solution Chatbot:
Is offered to people whilst they wait for their first appointment.
It is intended to help a client think about a problem or struggle they have, and then takes them through a 'brief solution focused coaching' conversation aiming to help them identify (and take) 1-2 small steps which they think might be helpful.

Things the Solution chatbot might say:

- Explores an issue or concerned the client may have whilst they wait for their first appointment
- Guides the client through a brief problem solving / solution focussed conversation once they have identified a problem
- Helps the client think of 1 or 2 actions they can take to overcome their problem

Prepared Client Chatbot:
Is offered to people whilst they wait for their first appointment.
It is designed to help a client think about and prepare for their first appointment, so that it can be as productive as possible.
There is the option to forward the result of the conversation to the therapist. The chatbot will ask about values, concerns and hopes for therapy

Things the Prepared Client chatbot might say:

- Asks the client to think about what would make their first therapy session as helpful as possible
- Asks the client what really matters to them in relation to their therapy
- Asks the client to think about what questions they have for their therapist

The pathway through the IAPT programme and the stages at which the chatbots are offered are shown in Figure 1. Figure 2 shows an example of what a client using one of the chatbots would see.

Figure 1 The IAPT pathway and chatbot placement

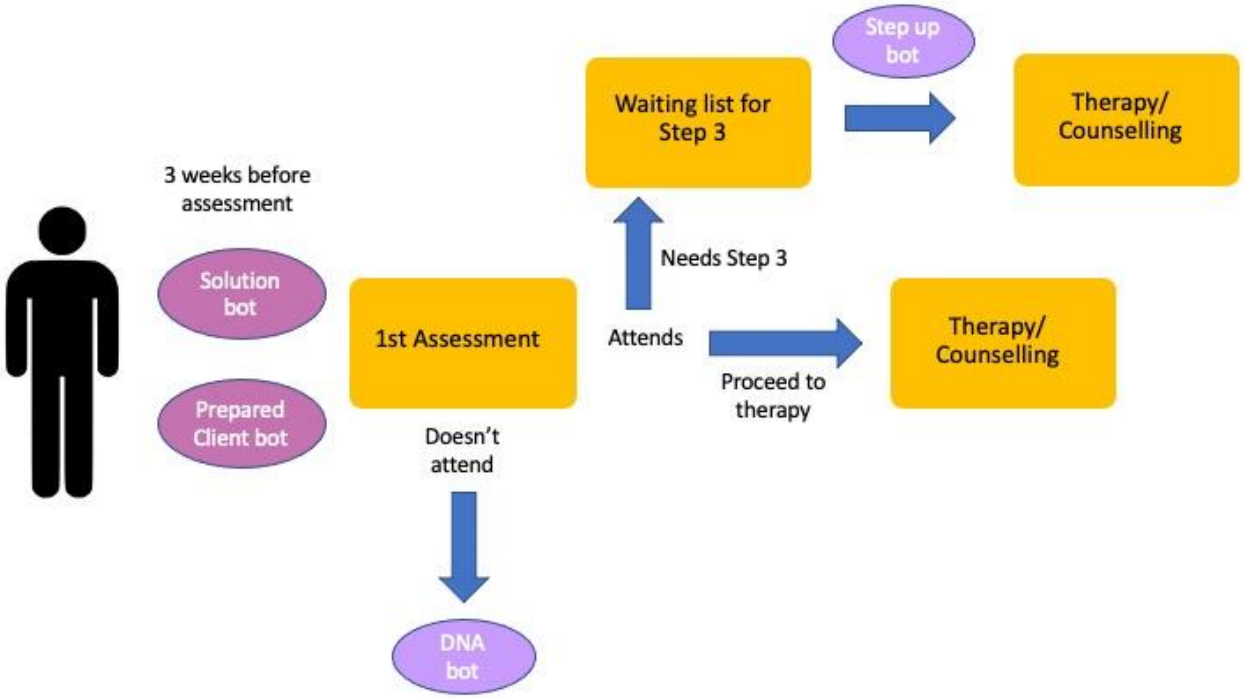
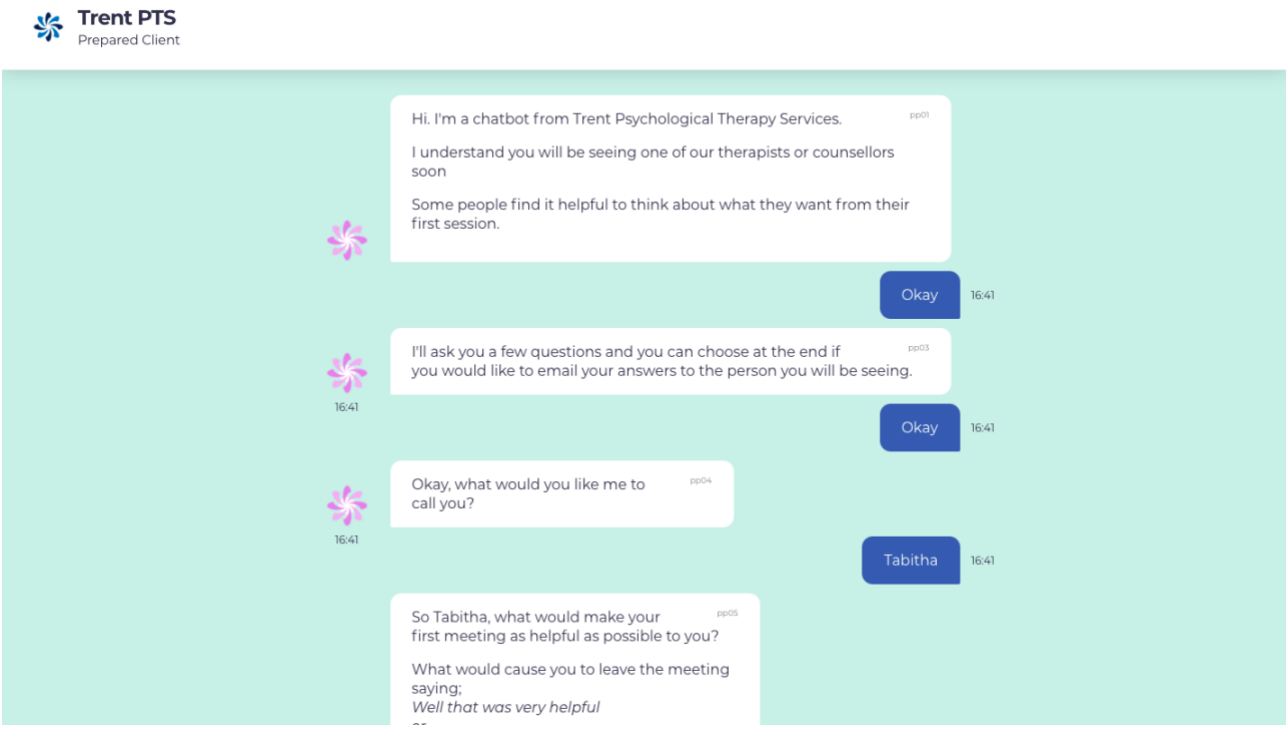


Figure 2 Example of chatbot



2.3 Sampling and Data Collection

Client Survey

A questionnaire was administered via the online platform Qualtrics to all clients who had used either the DNA or Step Up chatbots (n=418). In total 40 clients responded but 14 of these did not answer any questions reducing the final sample with data to 26 clients.

Due to the low uptake of the questionnaire the invitation that was sent to clients was changed to see if this would increase participation. In total 4 versions were used over the course of the evaluation, but all invitations elicited similar response rates.

The questionnaire was designed in consultation with the Project Steering Group (PSG) and consisted of 8 Likert scale questions regarding the client's satisfaction with the chatbot and 5 open ended questions to learn about the client's experience of the chatbot. The questionnaire was designed to be brief and to be completed in under 10 minutes to increase the uptake of participants. Therefore, the questionnaire did not include any demographic questions to reduce length and time taken to complete.

Client Feedback Data

Trent PTS were able to gather feedback from clients using the Solution and Prepared Client chatbots. At the end of the session the chatbots asked clients either a) what was it you found helpful? (Solution chatbot) or b) do you think this might help us offer a better service to our clients? (Prepared Client chatbot).

Clients were free to respond in any way they wanted. The responses (n=227) over a 4-week period were shared by Trent PTS and used for analysis. Over this 4-week period 123 responses were collected by the Solution chatbot and 104 from the Prepared Client chatbot.

Service Use Data

Data were provided by Trent PTS regarding their clients. This data consisted of:

- Demographics (e.g., age, diagnosis, ethnicity)
- Data regarding the number of appointments that had been attended
- Data regarding the number of DNAs
- Scores on psychometric scales (GAD-7 and PHQ-9) routinely administered by Trent PTS at the start and end of treatment.

Staff Focus Groups and Interviews

Six members of staff participated in an online focus group via MS Teams to discuss their views regarding the use of chatbots in the IAPT service. Additional focus groups were advertised but staff members declined to take part due to time commitments and only 1 more staff member could be recruited. Therefore, this staff member (psychological wellbeing practitioner) took part in an individual interview via Teams. The staff members consisted of counsellors, CBT therapists, psychological wellbeing practitioners and admin staff.

The topic for the focus group and interview was the same and was developed in consultation with the PSG. The questions were designed to learn what staff members currently working for Trent PTS thought about using chatbots in the IAPT service.

2.4 Data Analyses

Quantitative Data

The Likert scale data from the client survey were analysed using descriptive statistics. The service use data provided by Trent PTS were analysed using inferential statistics including Analysis of Variance (ANOVA), t tests and Chi square tests. All statistical data analyses were performed using IBM SPSS statistics (version 28).

Qualitative Data

Qualitative data from the focus group/interview with staff and qualitative data from open ended feedback questions were analysed using thematic analysis to identify common patterns and themes (Braun & Clarke, 2006).

2.5 Research Ethics

Ethical approval for the evaluation was obtained from Nottingham Trent University's Research Ethics Committee. All clients completing the survey via Qualtrics were given information regarding the evaluation and gave informed consent before commencing. Staff taking part in focus group/interview were emailed participant information sheets and consent forms. All secondary data that were shared by Trent PTS was in anonymised form, had no

identifiable information relating to individual clients and were shared in accordance with the Data Sharing Agreement.

3. Findings

3.1 Context

This section explores the data that relates to the context the chatbots are situated in and examines what the enablers and barriers are to implementing the chatbot.

DNA and Step Up Chatbots

The DNA and Step Up chatbots were originally implemented by Trent PTS in March 2021. Trent PTS reported there was a low uptake of these chatbots. Data showed that between March 2021 and August 2021 of the clients sent the text for the DNA chatbot**,** only **16.55%** went on to use the DNA chatbot**.** Of the clients sent the invitation text for the Step Up chatbot **24.55%** went on to use the Step Up chatbot.

Step up chatbot users (n =178) who used the chatbot between March 2021 and August 2021 had a mean age of **37.93 years.** The data showed that **70.22%** of users were female, **29.21%** were male and **0.56 %** did not have a gender recorded**.** The most common diagnosis was a Depressive Episode **(38.76%)** followed by Generalised Anxiety Disorder **(17.98%)**. When looking at the ethnicity of users it was found that **80.9%** were White, **6.18%** Asian, **5.06%** of Mixed Ethnicity, **3.37%** Black, **1.69%** Any other Ethnic Group and **2.81%** chose not to disclose their ethnicity.

DNA chatbot users (n =240) who used the chatbot between March 2021 and August 2021 had a mean age of **30.61 years**. It was found that **60.83%** of users were female and **39.17 %** were male. The most common diagnosis was a Depressive Episode **(43.14%)** followed by Generalised Anxiety Disorder **(19.61%)**. When looking at the ethnicity of users it was found that **80.83%** were White, **6.25%** Asian, **5.42%** of Mixed Ethnicity, **4.17%** Black, **1.67%** Any other Ethnic Group and **1.67%** chose not to disclose their ethnicity.

Solution and Prepared Client Chatbots

Following the low number of clients choosing to use the previous chatbots, two new chatbots were developed and in October 2021 the Solution and Prepared Client chatbots were launched.

Solution chatbot users (n =361) who used the chatbot between October 2021 and November 2021 had a mean age of **37.14 years.** It was found that **66.48%** of users were female, **32.96%** were male and **0.55 %** did not have a gender recorded**.** The most common diagnosis was a Depressive Episode **(43.07%)** followed by Generalised Anxiety Disorder **(23.76%)**. When looking at the ethnicity of users it was found that **85.87%** were White, **5.26%** Asian, **4.43%** of Mixed Ethnicity, **1.94%** Black, **0.55%** Any other Ethnic Group and **1.94%** chose not to disclose their ethnicity.

Prepared Client chatbot users (n =360) who used the chatbot between October 2021 and November 2021 had a mean age of **33.57 years.** The data showed that **63.33%** of users were female, **34.17%** were male and **2.5 %** did not have a gender recorded**.** The most common diagnosis was a Depressive Episode **(40.91%)** followed by Generalised Anxiety Disorder **(18.18%)**. When looking at the ethnicity of users it was found that **81.11%** were White, **6.11%** of Mixed Ethnicity**, 5.83%** Asian, **2.78%** Black, **1.67%** Any other Ethnic Group and **2.5%** chose not to disclose their ethnicity.

More detailed data collected over a 4 week period by the Solution and Prepared client chatbots were used to examine how many clients who started to use the chatbot went on to complete the full session. It was found that of the clients who started to use the Solution chatbot (n = 446) **31.17% went on to complete all questions** and finish the session. The most common place to exit the chatbot before completion was when being asked to name their problem with **49.78% of clients leaving the chat when asked to name the problem.**

Of clients who started a session with the Prepared Client chatbot (n = 473) **44.19% went on to complete all questions**. The most common place for clients to exit the chatbot before completion was when the chatbot started to ask questions with **22.83% of clients leaving when the chatbot said 'I'll ask a few questions.'**

Clients Who Chose Not to Use the Chatbot

Data were available for clients who were sent the text for the DNA and Step Up chatbots but did not go on to use it. In total 1211 clients did not respond to the text invitation of the DNA chatbot and 547 clients did not respond to the invitation for the Step Up chatbot. It was not possible to obtain this data within the timeframe of the evaluation for clients who had chosen not to use the Solution and Prepared Client chatbot. The data in Tables 2 and 3 shows a comparison between clients who responded to the text invitation and those who did not respond to the text and therefore did not use the chatbot (note this group of clients is distinct from those who were never offered the chatbot and acted as a control group).

Table 2 Comparison between clients who choose to use the DNA chatbot and clients who do not respond to the invitation

Variable	DNA chatbot users	Clients who opted not to use the DNA chatbot
Gender	60.83% female 39.17% male	62.81 % female 36.94 % male, not specified 0.25%
Age	Mean = 30.61 years (sd = 10.4)	Mean = 30.66 years (sd = 11.23)
Ethnicity	Asian = 5.3% Black = 2.8 % Mixed ethnicity = 5.2%	Asian = 4.79% Black = 2.64% Mixed ethnicity = 4.96%

	White = 85.9% Any other Ethnic group = 0.7%	White = 85.29% Any other Ethnic group = 0.5% Not specified = 1.82%
Area of Deprivation	No = 72.5% Yes = 27.5%	No = 72.91% Yes = 27.09%
Diagnosis	Most common Depressive episode = 43.14% Generalised anxiety disorder = 19.61%	Most common Depressive episode = 47.17% Generalised anxiety disorder =18.4%

Table 3 Comparison between clients who choose to use the Step Up chatbot and clients who do not respond to the invitation

Variable	Step up chatbot users	Clients who opted not to use the Step Up chatbot
Gender	70.22% female 29.21% male, 0.56% not specified	65.08 % female, 34.19 % male, 0.73% not specified
Age	Mean = 37.93 years (sd = 15.04)	Mean = 36.17 years (sd = 15.11)
Ethnicity	Asian = 6.18% Black = 3.37% Mixed ethnicity = 5.06% White = 80.9% Any other Ethnic group = 1.69% Not stated = 2.81%	Asian = 4.75% Black = 2.74% Mixed ethnicity = 3.84% White = 85.74% Any other Ethnic group = 1.28% Not stated = 1.65%
Area of Deprivation	No = 80.9% Yes = 19.1%	No = 82.2% Yes = 17.8%
Diagnosis	Most common Depressive episode = 38.76% Generalised anxiety disorder =17.98%	Most common Depressive episode = 43.88% Generalised anxiety disorder =16.82%

Summary

The data suggest that the DNA and Step Up chatbot had a low uptake with only 418 clients using the chatbots over 6 months. The Solution and Prepared Client chatbots were more popular with 721 clients using the chatbots in just 1 month.

When comparing clients who used the DNA and Step Up chatbot with clients who did not respond to the text invitation, no differences were found and the samples were similar.

Data extracted from the Solution and Prepared Client chatbots showed that out of the clients who started a conversation with the chatbot fewer than half continued through all the steps and finished the conversation. This suggests that not only is the initial engagement with clients to use the chatbot challenging their continued engagement remains similarly so.

3.2 Inputs

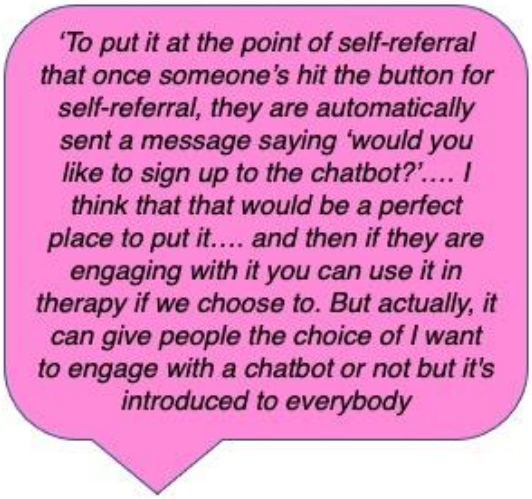
This section examines how the chatbot impacts on the IAPT service from staff perspective using data collected from staff members.

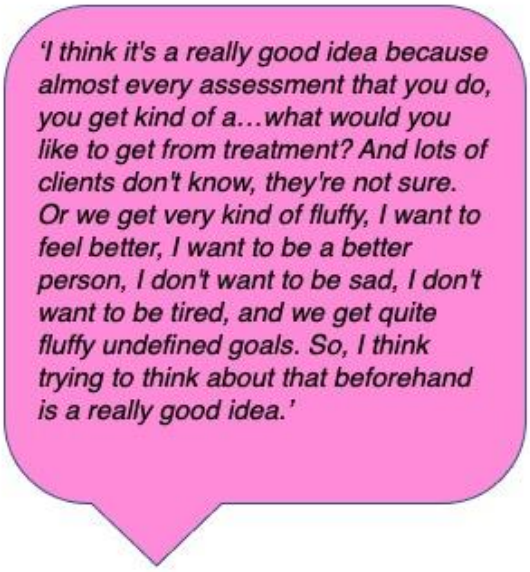
Focus Group Data

Qualitative data captured during the focus group and interview with staff members were analysed thematically to identify key areas that were thought to be important when considering the use of chatbots within the IAPT programme. The main themes identified are shown below with quotes from staff members who took part in the focus group used as examples.

When the chatbot is introduced to clients

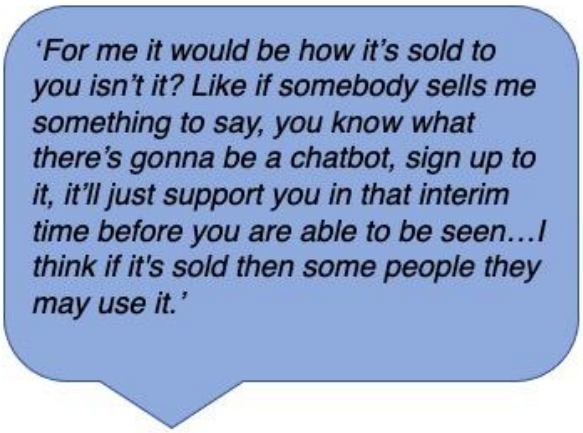
Staff felt it important that a chatbot is introduced at the right point in the IAPT pathway. It was felt that the chatbot should be introduced to everyone at the point of referral giving all clients the option of using the chatbot if they wanted to. Staff felt that using the chatbot to prepare for therapy would be a useful option.

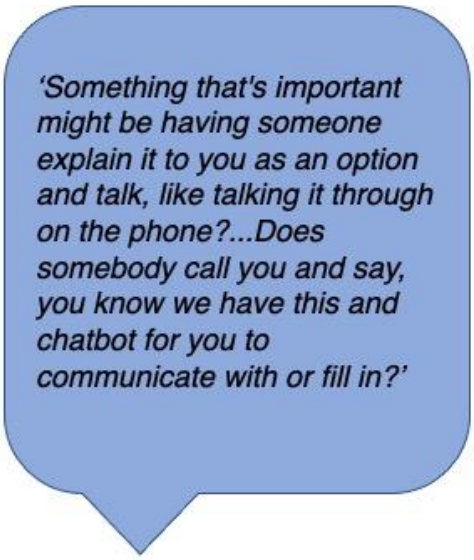




How the chatbot is sold to clients

The way in which the chatbot is introduced and 'sold' to clients was thought to be important. Staff thought that if the chatbot was promoted positively to clients they would be more likely to engage with it. Positive promotion was considered to include an introduction to the chatbot by a person rather than by a text message.





Which clients are offered the chatbot

The client group who are targeted by the chatbot was felt to be important. Staff felt some clients such as those who DNA would not be the right group to engage with a chatbot as they may lack commitment.

'My thing with people who DNA assessment is they're probably not likely to engage with a chatbot…Some people might but people who DNA assessments I wonder about their commitment to therapy. Are they then going to commit to engage with a chatbot? But I think it'd be really useful for people who are in therapy or waiting to go to therapy.'

"It's got to be something that… it's up to them if they want to use it or not and they know that it's if they want to use it or not. I think then it would be better cause the people that were using it were the people that wanted to use it and it wasn't something that was forced on them.'

What the chatbot is being used for

Staff considered that the chatbot would be good as part of therapy. Staff felt the chatbot could help facilitate clients' independence and lessen the chance of them becoming dependent on therapy. It was acknowledged that not all clients would want to engage with a chatbot and some could become frustrated with this.

'I think around CBT type stuff… so like reminders of what to do between sessions I thought that might be a good idea as long as the clients on board with it, sort of happy to engage with it.'

'It's just facilitating that independent management of symptoms rather than kind of just, you know, I'm feeling better 'cause I speak to somebody every other Tuesday. It's more about that taking on that responsibility of what can I do now, rather than creating a dependence on therapy'.

'I think in the context of therapy it feels quite limited because we're not really transactional like a bank we're more like in a relationship with people. So, I think it's got the potential to make some people really angry and frustrated, so that might not be great.'

'Like between sessions stuff, so anything that can help people engage with stuff between sessions or help to keep them motivated. You know if they are willing to give it a go. I think it would be helpful'

3.3 Outcomes

This section illustrates the findings that relate to the outcomes for both Trent PTS and its clients.

Client Survey

DNA and Step Up Chatbots

The data from the client survey (n=26) were analysed to understand if clients had a positive experience of using the chatbot. The open-ended questions on the survey were only completed by a small number of participants (n=9) meaning this part of the survey could not be analysed and only data from the Likert scale was examined.

Figure 3 shows the clients' responses to the 8 Likert scale questions regarding their use of either the DNA or Step Up chatbot.

Figure 3 Clients' responses to the questions

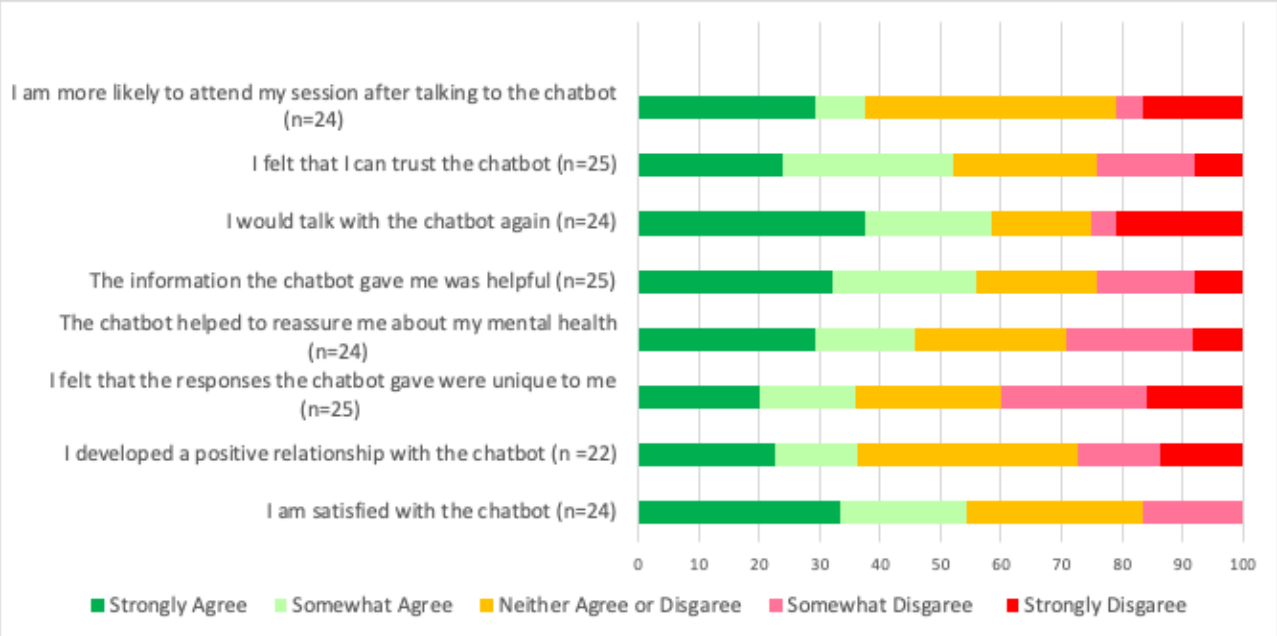


Table 4 Percentage of clients who gave a positive response (i.e., 'strongly agree' or 'somewhat agree')

Question	Percentage positive response
I am more likely to attend my session after talking to the chatbot	37.5%
I felt that I can trust the chatbot	52%

I would talk with the chatbot again	58.33%
The information the chatbot gave me was helpful	56%
The chatbot helped reassure me about my mental health	45.83%
I felt the responses the chatbot gave were unique to me	36%
I developed a positive relationship with the chatbot	36.36%
I am satisfied with the chatbot	54.17%

Summary

The data show that for 4 of the questions over half of clients responded positively saying that they **trusted the chatbot, would talk with it again, found it helpful** and were **satisfied with the chatbot.**

Client Feedback

The Solution and Prepared Client chatbot asked clients after completing their session if they had any feedback. Clients were free to respond in any way they wanted and the feedback was collected and analysed thematically to attain common themes that emerged across participants responses. Responses for both the Solution chatbot (n =123) and the Prepared Client chatbot (n = 104) were analysed separately.

Solution chatbot

Feedback collected from clients using the Solution chatbot was analysed thematically. When examining the responses overall and grouping them into either positive, negative, or neutral comments it was found that **50.37% of the responses were positive**, **28.15% were negative** and **21.48% of responses were neutral**

The most common themes are shown in Table 5. Overall, there were 6 main themes identified from the responses. From the table it can be seen that 4 of the themes related positively to the use of the chatbot and 2 of the themes identified concerns that clients had when using the chatbot.

Table 5 Main themes identified from client feedback on the Solution chatbot

Theme	Description	Example quotes

Things I can do to support myself	Clients expressed that the chatbot had given them the ability to be able to help themselves.	*"Suggestions on identifying things that I can work on while waiting for CBT.'* *'Being set an achievable target,'*
It made me think	Clients spoke of how the chatbot had caused them to reflect and think about their situation	*'Makes you think about things in a structured way.'* *'Admitting my issues to myself.'*
Being able to talk about my problem	Some clients said they liked having the opportunity to talk about things with the chatbot	*'Helped to talk through and unravel part of the problems I'm dealing with'* *'Not feeling alone with the problem.'*
Helps me calm down	Clients expressed how using the chatbot had made them feel calmer and helped to focus them	*'When you're feeling overwhelmed it's useful to have something impartial like this.'* *'It helped me calm down and breath for a second so I could think more rationally about the situation than emotionally like I was before.'*
I can't do it alone	Certain clients did not feel that the chatbot gave them the help they needed and felt they needed more guidance	*'I don't know how to fix my problem hence why I signed up to therapy.'* *'I can't solve any of my issues on my own so none of these questions helped me as I don't know the answer myself."*
How can I talk to a computer?	Some clients expressed their feelings of frustration at talking with a chatbot	*'I would prefer talking to someone rather [than] question time on a device.'* *'Not massively helpful and a bit impersonal.'*

Prepared Client Chatbot

The feedback collected from clients using the Prepared Client chatbot were analysed thematically. Overall, when looking at responses from clients and grouping them into either positive, negative or neutral thoughts regarding the chatbot it was found that **67.62% of the responses were positive**, **11.43% were negative** and **20.95% of comments were neutral**. In total, there were 3 main themes identified from the responses, as shown in Table 6.

Table 6 Main themes identified from client feedback on the Prepared Client chatbot

Theme	Description	Example quotes
Helpful	Clients found the chatbot helpful and were positive about its use	*'Great and insightful.'* *'I enjoyed this message chat and think it's helpful.'*
Reflection and preparation	Clients found it useful to be given the opportunity to prepare for their first session and valued the chatbot's role in this	*'Love the chatbot idea, some people may find it very impersonal but I think it has helped me be a little more prepared for my first session.'* *'I think it's a great way to start our sessions, so they have a quick insight into how I feel.'*
Anonymity	Some clients preferred not having to speak to someone face to face	*'This is a lot easier to answer as it can be quite hard to say these in front of a person.'* *'I think it might help some people who struggle to talk.. feels better if they can get it on paper.'*

Summary

The data shows that for both chatbots over half the clients gave positive responses when asked to give feedback. The Solution chatbot helped clients think clearly, think of ways to help themselves, reduced anxiety and gave clients an opportunity to talk about their problems. The analysis showed that clients using the Solution chatbot had some concerns around talking to a chatbot and some felt that they were being left alone without help. These issues were not apparent with the Prepared Client chatbot with most clients finding the Prepared Client chatbot positive and useful in helping prepare them for their first session. They also valued the anonymity this chatbot provided. Although the Prepared Client chatbot gives clients the option to forward the conversation to the therapist it may be that clients prefer initially saying personal things in an anonymous format rather than face to face the first time they meet the therapist.

Impact of Chatbots on Future Appointments, Cancellations, and DNAs

Step Up Chatbot

Data were analysed from the total number of clients who had used the Step Up chatbot up between March 2021 and August 2021 (n =178), and a control group of clients who were also on the waiting list for Step Up therapy but who were not offered the chatbot (n = 220)

An independent t test was conducted to identify any differences between the total number of attended appointments between clients who used the chatbot and those who did not. The results of the analysis showed that clients who had used the Step Up chatbot had significantly less total attended appointments than those who had not used the chatbot, $t(395.9) = 7.92$, $p<.001$. The findings are shown in Table 7.

Table 7 Mean number of total attended appointments for chatbot users and non-users

Group	Mean of attended appointments
Step up chatbot users	4.37 (sd =2.22)
Non chatbot users	6.33 (sd = 2.71)

The results highlight that on average chatbot users went on to have significantly fewer appointments when they started their course of therapy than those who had not used the chatbot.

To examine the outcomes of these groups, their scores for the GAD-7 and PHQ-9 were examined. These scales are administered by Trent PTS to clients at the beginning and end of therapy and therefore clients have 2 scores (First and Last). Higher scores on the GAD-7 signal a higher incidence of anxiety symptoms, therefore a reduction in scores means a reduction in anxiety. Higher scores on the PHQ-9 indicated a higher incidence of depressive symptoms, meaning a reduction in scores would suggest a reduction in depressive symptoms.

To examine any differences in these scores over the course of therapy for both chatbot users and non users mixed ANOVAs were conducted using Time (First; Last) and Chatbot user (Yes; No) as independent variables and GAD-7 and PHQ-9 scores as the dependent variables. A significant interaction was found between the variables for both GAD $F(1,390) = 33.93$, $p <.001$ and PHQ $F(1,390) = 59.58$, $p <.001$ scores suggesting that although all clients experienced a reduction in GAD and PHQ scores this reduction was significantly larger for clients who had not used the chatbot. This illustrated on Figures 4 and 5.

Figure 4 Changes in GAD scores for Step Up chatbot users

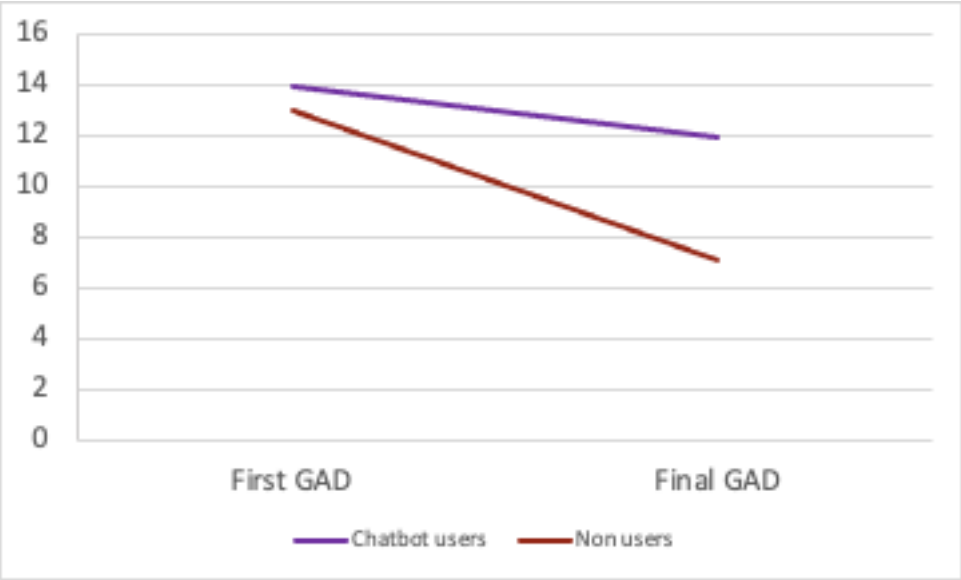
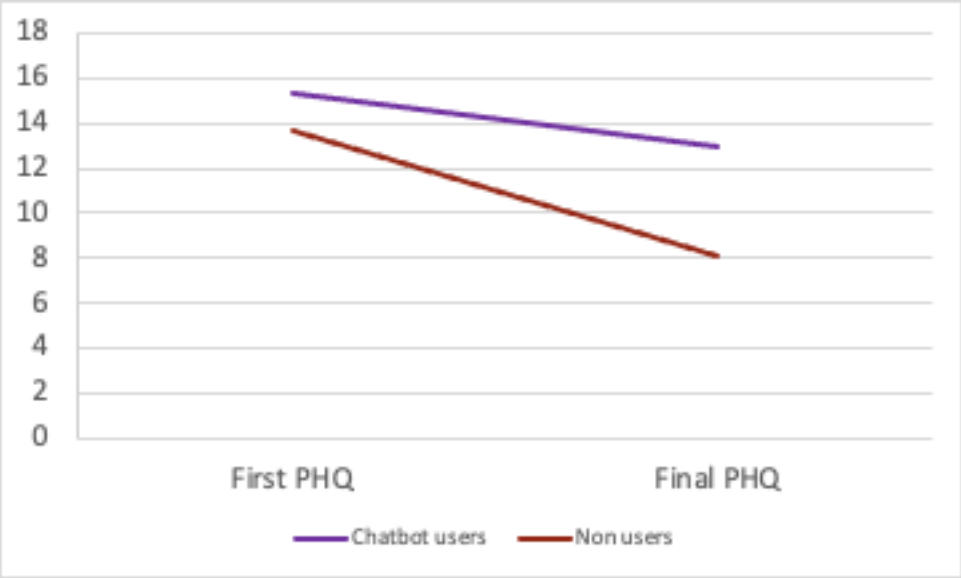


Figure 5 Changes in PHQ scores for Step Up chatbot users



The figures illustrate a significantly greater reduction in scores for clients who did not use the chatbot, this could be linked to the finding that these clients had more therapy sessions. However this needs further investigation and cost-effectiveness analysis to examine if the benefits of having less sessions outweigh the negatives.

DNA Chatbot

Data were **analysed** from the total number of clients who had used the DNA chatbot up between March 2021 and August 2021 (n = 240) and from clients in the same time period who had also not attended their assessment but who were not offered the chatbot (n = 458).

Independent t-tests were used to look at the difference in the amount of future DNAs and attended appointments between clients who used the chatbot and clients who did not. The results showed that clients who used the chatbot did not differ from those who had not used the chatbot in the amount of future DNAs, $t(696) = 0.78$, $p>.05$ with the number of future DNAs remaining similar between the groups. The analysis found that clients who used the chatbot had significantly fewer total attended appointments when compared to those who did not use the chatbot, $t(612.6) = 1.78$, $p=<.04$. The results are shown in Table 8.

Table 8 Mean numbers of total attended appointments and of total future DNAs for chatbot users and non-users

Group	Mean of attended appointments	Mean of future DNAs
DNA chatbot users	0.39 (sd =1.24)	1.16 (sd = 0.43)
Non chatbot users	0.59 (sd = 1.65)	1.14 (sd = 0.38)

The findings show that the number of total appointments was significantly lower in those who used the chatbot. From the data it is evident that the total number of attended appointments is relatively low for both groups suggesting that clients who DNA do not typically go on to attend a high number of appointments. This suggests that those who DNA are a critical group for intervention as they are not going on to complete therapy. The results also show that on average clients went on to repeat DNA for another appointment and this could not be prevented by using the chatbot.

To further examine the profile of clients who DNA their assessment the main characteristics of a sample of clients who had DNA'd (n = 1908) were compared with a control sample of clients who did not DNA (n = 1907) to ascertain if there are any differences in key variables. T tests and Chi square tests were conducted to identify these differences between groups. The results are shown in Table 9 where it can be seen that age and deprivation background (whether clients come from an area of deprivation) are significantly different between the groups.

Table 9 Comparison between DNA and non-DNA clients

Variable	DNA clients	Control group clients	Significance level
Gender	63.3% female 36.7% male	64.6 % female 35.4 % male	p = >.05

Age	Mean = 30.1 years (sd = 11.29)	Mean = 35.86 years (sd 14.96)	***p* = <.001**
Ethnicity	Asian = 5.3% Black = 2.8 % Mixed ethnicity = 5.2% White = 85.9% Any other Ethnic group = 0.7%	Asian = 5.8% Black = 1.9% Mixed ethnicity = 4.1% White = 87.3% Any other Ethnic group = 0.9%	*p* = >.05
Area of Deprivation	No = 72.7% Yes = 27.3%	No = 82.2% Yes = 17.8%	***p* = <.001**
Diagnosis	Most common Depressive episode = 46.9% Generalised anxiety disorder =21.6% Mixed anxiety and depressive disorder = 8.2%	Most common Depressive episode = 46.3% Generalised anxiety disorder =18.2% Post traumatic stress disorder = 8%	*p* = >.05

Solution and Prepared Client Chatbots

Data were analysed from the 2 new chatbots to determine if the chatbots had any effect on DNAs and cancellation of appointments. A 1 way ANOVA comparing the amount of DNAs across clients who had used the Solution chatbot (n = 361), clients who used the Prepared Client chatbot (n =360) and a control group (n = 721) of clients who hadn't used a chatbot was conducted. This showed that the amount of DNAs were significantly different across groups, $F(2,1441) = 9.98$, $p < .001$. Post hoc analysis using Tukey's HSD showed that DNAs were significantly lower for clients using the Solution chatbot compared to clients who used the Prepared Client chatbot ($p < .001$) but were not significantly different from the control group ($p > .05$). DNAs were significantly higher for clients using the Prepared Client chatbot compared to the control group ($p < .001$).

A 1-way ANOVA was also conducted to compare the amount of cancellations across the three groups. The results showed the amount of cancellations were significantly different across the groups $F(2,1441) = 6.03$, $p = 002$. Post hoc analysis using Tukey's HSD showed that clients using the Solution chatbot had significantly fewer cancellations than clients using the Prepared Client chatbot ($p = .002$) and clients in the control group ($p = .02$). There was no significant difference between clients using the Prepared Client chatbot and the control group ($p > .05$).

The findings are shown in Table 10 and Figures 6 and 7 where it is shown that the clients using the Solution chatbot had fewer DNAs and cancellations.

Table 10 Mean numbers of total DNAs and total cancellations for clients using Solution and Prepared Client chatbots

Group	Mean of total DNAs	Mean of total cancellations
Solution chatbot users	0.29 (sd = 0.54)	0.11(sd = 0.31)
Prepared Client chatbot users	0.49 (sd =0.73)	0.21 (sd = 0.41)
Non chatbot users (control group)	0.36 (sd = 0.57)	0.18 (sd = 0.39)

Figure 6 Mean number of DNAs across the Solution and Prepared Client chatbots
(Error bars represent the 95% confidence intervals)

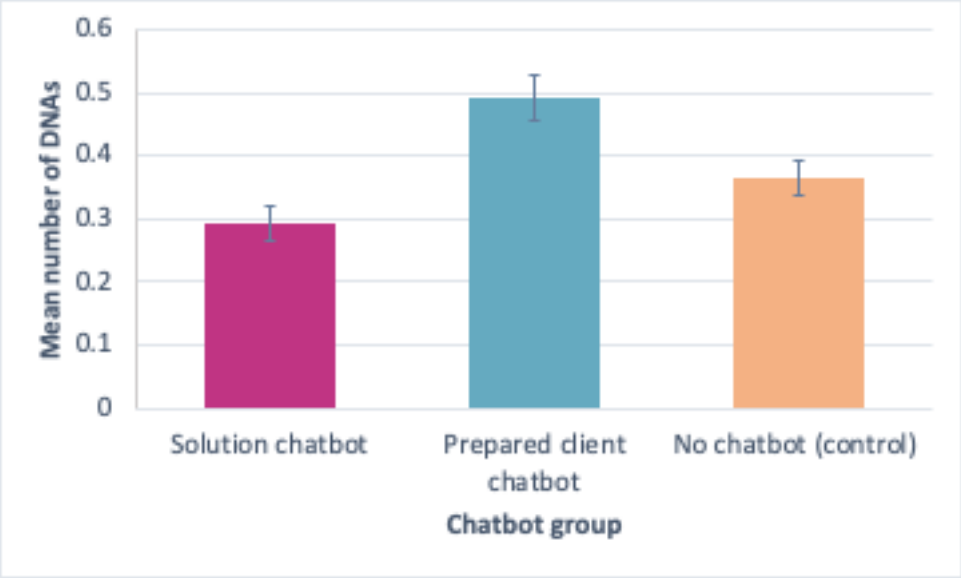
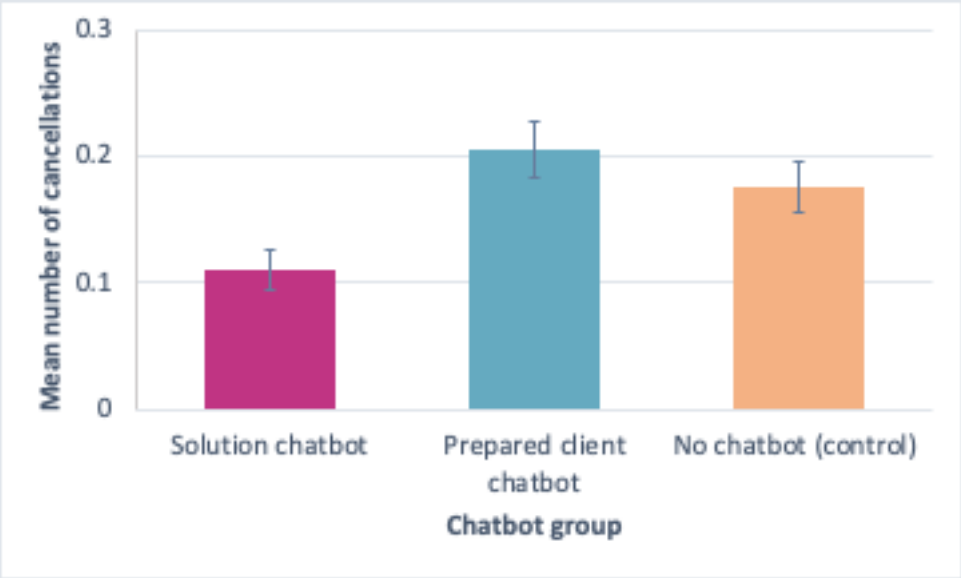


Figure 7 Mean number of cancellations across Solution and Prepared Client chatbots
(Error bars represent the 95% confidence intervals)



Summary

Clients who had used the Step Up chatbot had on average fewer appointments than those who did not use the chatbot. Total numbers of attended appointments were lower for clients who DNA indicating that this group are less likely to complete a full course of therapy. The DNA chatbot was not found to affect further DNAs in clients. When looking at the profile of clients who DNA the data suggests age and coming from an area of deprivation, are significant factors.

Data from the new chatbots implemented showed that Solution chatbot was more effective at reducing DNAs and cancellations compared to the Prepared Client chatbot. Cancellations were significantly lower for clients using the Solution chatbot compared to clients who did not use a chatbot and clients who used the Prepared Client chatbot. DNAs for clients who used the Solution chatbot were not significantly different to clients who did not use a chatbot but were significantly lower than clients who used the Prepared Client chatbot.

4. Conclusions

The findings suggest that feedback was largely positive from clients who had used the chatbots showing the potential of using chatbots within the IAPT programme. When the new chatbots were launched by Trent PTS the uptake was higher and client feedback remained positive suggesting these chatbots were more effective at engaging clients. This was thought to be because the chatbots target all clients and include content that can be related to therapy.

The findings show that when implementing a chatbot the target client group needs to be given some consideration. Trent PTS found that clients who DNA their assessment were reluctant to engage with a chatbot suggesting that this is not an effective way of targeting this particular sub-group. Trent PTS had more success in engaging clients when the new chatbots were launched which suggests uptake of the chatbot is increased when it's offered to all clients before assessment. Staff in focus groups also felt that clients who DNA were not the ideal group to target as their motivation to engage generally maybe low. Instead, a chatbot should be offered to all clients who could then choose if they wanted to use it. The data suggests that even for clients who did engage with the DNA chatbot it did not prevent further DNAs, also highlighting that a chatbot is not effective at reducing this.

The Solution and Prepared Client chatbots gave clients the opportunity to discuss a problem or prepare for their therapy. Feedback from clients indicated that this was popular suggesting clients find it useful when the chatbot is linked to the therapy process. The findings showed that for these chatbots most clients gave positive responses when asked to give feedback. It was shown that the Solution chatbot helped clients think clearly, think of ways to help themselves, reduced anxiety and gave clients an opportunity to talk about their problems. Clients were positive about the role of the Prepared Client chatbot in preparing them for therapy.

The findings suggested that the Solution chatbot has the potential to reduce non-attendance at appointments. It was found that the Solution chatbot seemed to be more effective at reducing DNAs and cancellation of appointments in comparison to the Prepared Client chatbot, this could be due to the different content of the chatbots. However more research is needed to explore this further and to establish what it is about the Solution chatbot that encourages clients to attend their appointments.

Despite the success of the Solution chatbot at reducing DNAs and cancellations the analysis showed that some clients using the Solution chatbot had some hesitance around talking to a chatbot. Feedback from the Prepared Client chatbot showed that clients did not raise these concerns when using this chatbot. This could be due to the specific topics discussed; the Solution chatbot asks clients to think about a problem and discuss how they can tackle it implying that the conversation may cover more sensitive or challenging topics than with the Prepared Client chatbot. It is also important to note that the chatbots asked different questions when asking the clients for feedback meaning differences in responses could be due to this.

The data from the Step Up chatbot showed that clients who had used this had fewer appointments. This finding would need to be explored further. It is possible that another unaccounted-for variable is confounding the results as the Step Up chatbot is not used during therapy by the clients so should not necessarily lead to them taking fewer sessions.

From the secondary data provided by Trent PTS on numbers of attended appointments, cancellations and DNAs, it is not possible to tell how much the client interacted with the chatbot of if they fully completed the session. More detailed data extracted from the chatbots showed that a proportion of clients will leave the chatbot before the conversation has ended. It is likely that the extent to which a client engaged with the chatbot would have affected the data.

The evaluation was also limited by the challenges around data collection. The client survey was only completed by a very small number of clients and therefore is not representative of all clients using these chatbots. However, client feedback collected by the chatbot was completed by a much larger sample and therefore gave a greater insight into clients' views of the chatbot. In addition, client feedback was only completed by clients who reached the end of the chatbot which is beneficial as it means that this data was only collected from clients who had fully engaged with the chatbot.

The data available for the evaluation did not allow us to establish the cost-effectiveness of the chatbots due to the complexity of achieving this. The solution chatbot showed the potential of reducing DNAs and cancellations which could result in cost savings over time but this would need to be examined over a longer time period than the 4-week period used here. Any cost savings also need to be offset against the costs taken to implement a chatbot, which means that it could be several months before a chatbot becomes cost-efficient.

The use of chatbots in psychological therapy and mental health support is a new and still developing area. This evaluation suggests there is potential for chatbots to be used within the IAPT service when they are directly related to the therapy process and when both clients and therapists can see the benefits of clients interacting with them. Future research is still needed to examine how the current chatbots being used by Trent PTS could evolve further and be fully integrated into therapy and the IAPT journey while continuing to improve outcomes for clients.

Recommendations

- Offer the chatbot to all clients before assessment. Clients who DNA their assessment were reluctant to engage with a chatbot, suggesting that this is not an effective way of solving this problem. There was more success in engaging clients when the new chatbots were launched and offered to all clients before assessment.

- Conduct further research to understand why the Solution chatbot is more successful than the Prepared Client chatbot at reducing DNAs and cancellations. The Prepared Client chatbot received positive feedback from clients despite not reducing DNAs or cancellations. This highlights that how a client interacts and feels about a chatbot will not necessarily impact on their behaviour. If possible, a model of offering clients both the Solution and Prepared Client chatbots should be trailed to assess how effective the chatbots are when combined. Or different chatbots at different stages could be offered to clients to assess how effective this is.

- Conduct an evaluation to compare individual chatbots at different times with the combined chatbot to test the hypothesis that engagement with a chatbot that promotes thinking about engaging with therapy (cognition) and practising solving a problem (behaviour), is more effective and met with greater satisfaction than when these two aspects of therapy are presented separately by the current chatbots.

- Conduct follow-up research over a longer time period that fully examines the cost effectiveness of using chatbots. This should be done using multiple indices of cost-effectiveness to individuals and services. For example, over time do the costs saved

as result of reducing non-attendance at appointments offset the costs of developing and/or administering a chatbot? Can any other cost savings be established such as the costs of less or more effective therapy sessions that could occur as a result of using chatbots?

- Use the chatbot alongside therapy and integrate it more closely with the therapy process. Make sure that clients understand how the chatbot and therapy are linked together and support each other. The chatbot should not be perceived as an add-on, a burden, a time consuming or irrelevant activity. Rather, it should be seen as essential to the therapy programme.

- Chatbot scripts should be considered in terms of how questions are asked (how direct it is, what the expectations are of the client, how prepared the client is to engage at this level with a chatbot) to try and maximise the number of clients who complete a full session with the chatbot.

- How the chatbot is introduced is important. A focus on explaining the purpose of the chatbot to the clients and presenting the chatbot as a tool to aid therapy.

- Continue to include staff members and clients in future development and design of the chatbots to gain their input and experience.

- Finally, conduct follow-up research that looks are the mechanisms of change for different groups of clients and at different stages of the process. The present findings have given us a glimpse of the possible reasons for the effectiveness of the chatbot, and the conditions for that, but additional in-depth research can further support the development of tailored efficient and effective chatbots.

5. Acknowledgments

The NTU research team would like to thank the following for their contributions and for the time they have given during the evaluation:

- Dean Repper, Clinical Services Director at Trent PTS, for his coordination of the data collection and his continued valuable input throughout the project, without his support the evaluation would not have been possible.

- Virtual Health Labs for designing, producing, and testing the chatbots, and specifically Dr Tim Anstiss for supporting the evaluation project.

- The Project Steering Group for their contributions during meetings and overall support throughout the evaluation.

- Staff at Trent PTS who took part in the focus group and interview and shared their views and experiences, and staff who were involved in the distribution of the client survey.

- Clients using the chatbots who gave feedback and shared their insights on using the chatbots helping us to understand how the chatbot impacts on their experience.

6. References

Abd-Alrazaq, A., Alajlani, M., Ali, N., Denecke, K., Bewick, B., & Househ, M. (2021). Perceptions and Opinions of Patients About Mental Health Chatbots: Scoping Review. *Journal of Medical Internet Research, 23*(1): e17828

Bailey, D., De Motte, C., Nomikos, L., & Mutale, G. (2020). *Evaluation of the You Know Your Mind Project*. Research report for external body. Nottingham: Nottingham Trent University.

Bailey, D., Holland, D., & Mutale, G. (2017). *Evaluation of the Social Care Role in Integrated Primary Care Teams for Older Adults who have Complex Needs in Nottinghamshire*. Research report for external body. Nottingham: Nottingham Trent University.

Bailey, D., & Kerlin, L. (2016). Can health trainers make a difference with difficult-to-engage clients? A multisite case study. *Health Promotion Practice, 16(*5), 756–64.

Bailey, D., & Mutale, G. J. (2020). Can a case lead approach deliver the "craft and graft" of integration? *Journal of Integrated Care, 29*(1), 72-84

Bell, S., Wood, C., & Sarkar, A. (2019). Perceptions of chatbots in therapy. CHI'19 Extended Abstracts, May 4th–9th 2019, Glasgow, Scotland UK

Biron, C., & Karanika-Murray, M. (2014). Process evaluation for organizational stress and well-being interventions: Implications for theory, method, and practice. *International Journal of Stress Management, 21*(1), 85–111

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology, 3*(2), 77–101

Fitzpatrick, K. K., Darcy, A., & Vierhile, M. (2017). Delivering Cognitive Behavior Therapy to Young Adults With Symptoms of Depression and Anxiety Using a Fully Automated Conversational Agent (Woebot): A Randomized Controlled Trial. *JMIR Mental health*, *4*(2), e19.

Karanika-Murray, M., Biron, C., & Saksvik, P.Ø. (2016). Guest editorial: Organizational health interventions: Advances in evaluation methodology. *Stress & Health, 32*(4), 255-257.

Kirkpatrick, D. L. (1994). *Evaluating training programs: the four levels.* San Francisco: Berrett-Koehler Pawson R, Tilley N. Realistic evaluation. 2019. Available from: https://www.communitymatters.com.au/RE_chapter.pdf.

Kretzschmar, K., Tyroll, H., Pavarini, G., Manzini, A., Singh, I., & NeurOx Young People's Advisory Group (2019). Can Your Phone Be Your Therapist? Young People's Ethical

Perspectives on the Use of Fully Automated Conversational Agents (Chatbots) in Mental Health Support. *Biomedical Informatics Insights*, *11*, 1178222619829083

Laranjo, L., Dunn, A. G., Tong, H. L., Kocaballi, A. B., Chen, J., Bashir, R., Surian, D., Gallego, B., Magrabi, F., Lau, A., & Coiera, E. (2018). Conversational agents in healthcare: a systematic review. *Journal of the American Medical Informatics Association : JAMIA*, *25*(9), 1248–1258.

Lucas, G. M., Gratch, J., King, A., & Morency, L.-P. (2014). It's only a computer: Virtual humans increase willingness to disclose. *Computers in Human Behavior, 37,* 94–100

Miner, A. S., Milstein, A., Schueller, S., Hegde, R., Mangurian, C., & Linos, E. (2016). Smartphone-Based Conversational Agents and Responses to Questions About Mental Health, Interpersonal Violence, and Physical Health. *JAMA internal medicine*, *176*(5), 619–625.

Mutale, G. J., De Motte, C., & Bailey, D. (2020). Using a mixed methods approach to examine the effectiveness of inpatient rehabilitation services following a programme of planned bed closures. *Journal of Psychosocial Rehabilitation and Mental Health*, *7,* 55-65

NHS (2021). Adult Improving Access to Psychological Therapies programme. Available from https://www.england.nhs.uk/mental-health/adults/iapt/

NHS (2019). *NHS Mental Health Implementation Plan 2019/20 – 2023/24.* Available from: https://www.longtermplan.nhs.uk/wp-content/uploads/2019/07/nhs-mental-health-implementation-plan-2019-20-2023-24.pdf

Vaidyam, A. N., Wisniewski, H., Halamka, J. D., Kashavan, M. S., & Torous, J. B. (2019). Chatbots and Conversational Agents in Mental Health: A Review of the Psychiatric Landscape. *Canadian Journal of Psychiatry*, *64*(7), 456–464

Ward, J., & Bailey, D. (2015). How far can a short leadership and management programme address the challenges for first line social work managers? An evaluation of one of the skills for care leadership and management demonstration sites. *Practice Social Work in Action, 28,* 81–303.

Warr, P., Bird, M., & Rackham N. (1970). *Evaluation of management training*. London: Gower Press.