

Original Paper

Acceptability of a Pain History Assessment and Education Chatbot (Dolores) Across Age Groups in Populations With Chronic Pain: Development and Pilot Testing

Nicole Emma Andrews^{1,2,3,4}, BOccThy (Hons), PhD; David Ireland⁵, BE, PhD; Pranavie Vijayakumar^{5,6}, BBioMed (Hons); Lyza Burvill⁷, BPhy (Hons); Elizabeth Hay⁷, BPhy (Hons); Daria Westerman⁸, MOccThy; Tanya Rose⁷, BSpPath (Hons), PhD; Mikaela Schlumpf⁸, BOccThy; Jenny Strong^{2,3,7}, BOccThy, MOccThy, PhD; Andrew Claus^{2,7}, BPhy (Hons), MPhy, PhD

¹RECOVER Injury Research Centre, The University of Queensland, Herston, Australia

²Tess Cramond Pain and Research Centre, The Royal Brisbane and Women's Hospital, Metro North Hospital and Health Service, Herston, Australia

³The Occupational Therapy Department, The Royal Brisbane and Women's Hospital, Metro North Hospital and Health Service, Herston, Australia

⁴Surgical Treatment and Rehabilitation Service (STARS) Education and Research Alliance, The University of Queensland and Metro North Health, Herston, Australia

⁵Australian eHealth Research Centre, The Commonwealth Scientific and Industrial Research Organisation, Herston, Australia

⁶The Walter and Eliza Hall Institute of Medical Research, Melbourne, Victoria, Australia

⁷School of Health and Rehabilitation Sciences, The University of Queensland, St Lucia, Australia

⁸Queensland Interdisciplinary Paediatric Persistent Pain Service, Queensland Children's Hospital, South Brisbane, Australia

Corresponding Author:

Nicole Emma Andrews, BOccThy (Hons), PhD

RECOVER Injury Research Centre

The University of Queensland

Level 7, Surgical Treatment and Rehabilitation Service (STARS)

296 Herston Rd

Herston, 4029

Australia

Phone: 61 418762617

Email: n.andrews@uq.edu.au

Abstract

Background: The delivery of education on pain neuroscience and the evidence for different treatment approaches has become a key component of contemporary persistent pain management. Chatbots, or more formally conversation agents, are increasingly being used in health care settings due to their versatility in providing interactive and individualized approaches to both capture and deliver information. Research focused on the acceptability of diverse chatbot formats can assist in developing a better understanding of the educational needs of target populations.

Objective: This study aims to detail the development and initial pilot testing of a multimodality pain education chatbot (Dolores) that can be used across different age groups and investigate whether acceptability and feedback were comparable across age groups following pilot testing.

Methods: Following an initial design phase involving software engineers (n=2) and expert clinicians (n=6), a total of 60 individuals with chronic pain who attended an outpatient clinic at 1 of 2 pain centers in Australia were recruited for pilot testing. The 60 individuals consisted of 20 (33%) adolescents (aged 10-18 years), 20 (33%) young adults (aged 19-35 years), and 20 (33%) adults (aged >35 years) with persistent pain. Participants spent 20 to 30 minutes completing interactive chatbot activities that enabled the Dolores app to gather a pain history and provide education about pain and pain treatments. After the chatbot activities, participants completed a custom-made feedback questionnaire measuring the acceptability constructs pertaining to health education chatbots. To determine the effect of age group on the acceptability ratings and feedback provided, a series of binomial logistic regression models and cumulative odds ordinal logistic regression models with proportional odds were generated.

Results: Overall, acceptability was high for the following constructs: engagement, perceived value, usability, accuracy, responsiveness, adoption intention, esthetics, and overall quality. The effect of age group on all acceptability ratings was small

and not statistically significant. An analysis of open-ended question responses revealed that major frustrations with the app were related to Dolores' speech, which was explored further through a comparative analysis. With respect to providing negative feedback about Dolores' speech, a logistic regression model showed that the effect of age group was statistically significant ($\chi^2=11.7$; $P=.003$) and explained 27.1% of the variance (Nagelkerke R^2). Adults and young adults were less likely to comment on Dolores' speech compared with adolescent participants (odds ratio 0.20, 95% CI 0.05-0.84 and odds ratio 0.05, 95% CI 0.01-0.43, respectively). Comments were related to both speech rate (too slow) and quality (unpleasant and robotic).

Conclusions: This study provides support for the acceptability of pain history and education chatbots across different age groups. Chatbot acceptability for adolescent cohorts may be improved by enabling the self-selection of speech characteristics such as rate and personable tone.

(*JMIR Form Res* 2023;7:e47267) doi: [10.2196/47267](https://doi.org/10.2196/47267)

KEYWORDS

chronic pain; education; neurophysiology; neuroscience; conversation agent; chatbot; age; young adult; adolescence; adolescent; pain; patient education; usability; acceptability; mobile health; mHealth; mobile app; health app; youth; mobile phone

Introduction

Background

Although pain is commonly thought of as a symptom resulting from injury or disease, chronic or persistent pain is considered a disease in its own right [1]. The disease presents as a complex, common, and costly health problem [2]. The Global Burden of Disease 2016 study found that pain-related diseases, lower back pain, and migraine are the leading causes of disability and disease burden globally [3]. Pain is also the main reason why people seek medical care, with 3 of the top 10 reasons being treatment for osteoarthritis, back pain, or headaches [4]. Chronic pain is common across the life span, with high rates of recurrent headaches and abdominal pain reported in pediatric populations [5].

Most approaches to the management of chronic pain incorporate education about pain neuroscience and the evidence for different clinical interventions such as surgery, pain relief medications, and nonpharmacological approaches [6-8]. Used by both adult and pediatric populations, the education aims to shift one's conceptualization of pain from that of an indicator of tissue damage or disease to a central nervous system output designed to protect the body [9]. Thus, pain education is a clinical tool that empowers people to self-manage their condition through a better understanding of pain and uptake of evidence-based physical, psychological, and social treatment strategies [10,11].

Recent systematic reviews have revealed that pain education can facilitate the ability to cope with persistent pain by decreasing pain catastrophizing and kinesiophobia [7,11], and when combined with exercise therapy, it can result in greater short-term improvements in pain and disability than exercise alone in adults [7]. However, as most research has focused on educating adults with chronic pain, further empirical investigations with pediatric pain populations are required to effectively tailor and implement pain education [8,12].

To facilitate the processing of complex and novel pain education information in pediatric populations, it has been suggested that multiple modalities may be needed such as video clips and animated mobile phone apps [8]. Mobile phone apps might be particularly useful for digital native adolescents. However, some

studies have found low levels of engagement for pediatric pain self-management apps, suggesting that further research is needed to understand how adoption can be enhanced [13,14]. In an adult cohort, allowing people with chronic pain to be able to tell their own story through a comprehensive assessment has been identified as a key component in enhancing the experience of pain education [11].

Chatbots, or more formally conversation agents, are increasingly being used in health care settings to provide education [15]. Chatbots are computer systems that enable interactive 2-way communication through a variety of modalities including videos, images, and written or spoken language [15,16]. The versatility of these communication modalities enhances the potential application of educational chatbots in health care settings across a variety of target populations ranging from young children to older adults [15]. Chatbots are more individualized than leaflets and can be easier to navigate than websites. A survey of 100 physicians acknowledged that although chatbots cannot effectively address the needs of all patients, they can play an important role in supporting and motivating them, resulting in potential cost savings and improved patient outcomes [17].

However, there is a paucity of research on the development, automation, adoption, and acceptability of chatbots from the perspective of patients in the health field [18]. Two available chatbots were evaluated in cohorts consisting of adults with chronic pain, and both evaluations revealed high levels of user acceptability and engagement [19,20]. No known study has used a chatbot in a pediatric pain cohort. There is a need for further research that is focused on the acceptability of diverse chatbot formats, resulting in a better understanding of the needs and preferences of different age groups [15].

Objectives

The aims of this study were (1) to detail the development of a multimodality pain education chatbot (Dolores); (2) to conduct initial pilot testing with adolescents, young adults, and adults with chronic pain; and (3) to investigate whether acceptability and feedback were comparable across age groups following initial pilot testing. It was hypothesized that acceptability would be comparable across age groups based on a systematic review, indicating that age does not have a consistent effect on health information technology acceptance [21].

Methods

Chatbot Development and Framework

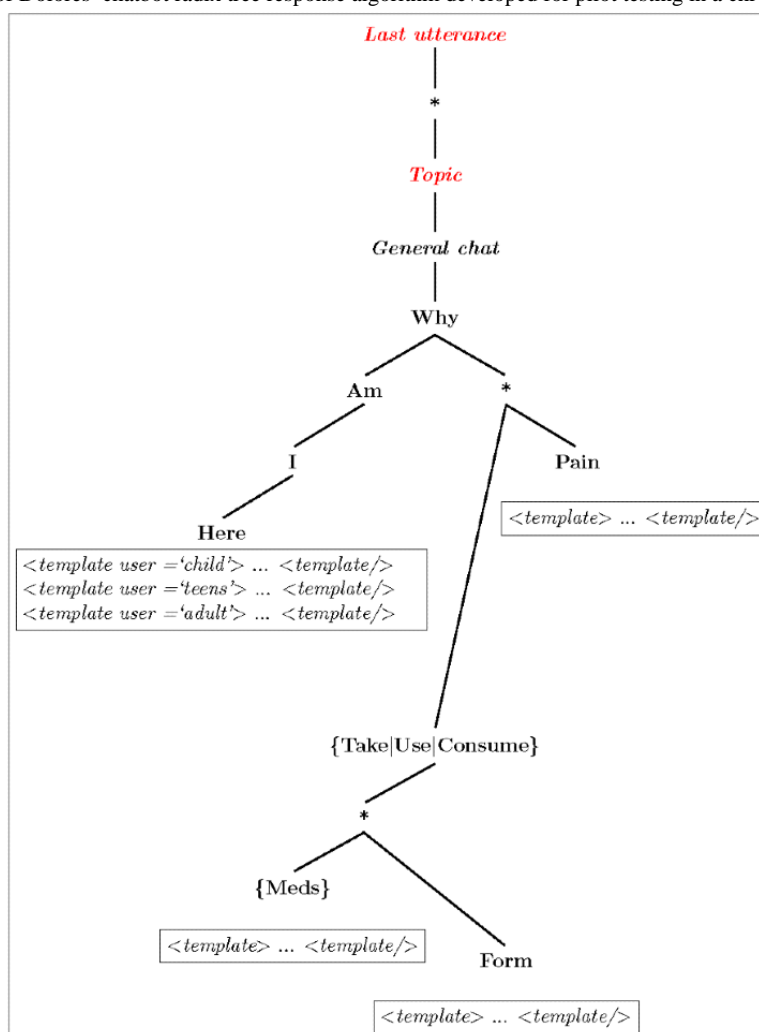
Dolores was designed to be embedded in the Pain ROADMAP monitoring app [22,23]. The idea for Dolores resulted from a working group with researchers and clinicians who discussed how to adapt Pain ROADMAP to a pediatric setting. The chatbot framework used to build Dolores is an in-house framework developed at the Commonwealth Scientific and Industrial Research Organisation (CSIRO) in Australia. This framework has been used in other health-related chatbots, including Harlie for autism spectrum disorder and Parkinson disease [24,25] and Edna for genetic counseling [26]. The response engine consists of a case-based reasoner that operates based on the structure (syntax) of human communication. The underlying data structure and response algorithm is a radix tree [24], where the nodes are the possible human-uttered words (or wildcard matching syntax). This data structure provides inherent compression of similar utterances and fast lookup times. A response is produced when the traversal of the radix tree structure finds a *template* response at the end of the utterance.

Figure 1 provides an example of this process. The radix tree is first created by parsing markup files that contain the patterns (possible human utterances) and corresponding templates that provide instructions on how to respond. The markup language used was based on XML. An example is shown in Textbox 1, which when parsed, forms a part of the tree shown in Figure 1.

The pattern can include wildcards (eg, *: ≥1 words), optional words, sets of words (ie, synonyms), and forbidden words. A constructed radix tree is referred to as a “brain.” The example given in Figure 1 would provide matched responses to utterances such as “Why do I feel pain?” and “Why do I need to take my meds?”

Auxiliary algorithms are also used including sentiment analysis that identifies the probability of negative sentiments and logic reasoning. Multiple responses are encoded, with each specific response chosen based on the sentiment of the last utterance and overall conversation [27]. When Dolores does not have a match in her brain, sentiment analysis is used to determine whether to refer the user to a human or request more information.

Figure 1. Illustrative example of Dolores’ chatbot radix tree response algorithm developed for pilot testing in a chronic pain population.



Textbox 1. Example of markup language for responding to utterances of the form “Why do I have pain?” “Why do I feel pain?” etc.

```
<topic name = “general-chat”>
<category>
<pattern> WHY * PAIN </pattern>
<template>
    Pain is a defense mechanism in the brain.
</template>
</category>
</topic>
```

The design and development of Dolores’ brain was a collaborative process among a group of 6 expert clinicians: 4 occupational therapists, 1 physiotherapist, and 1 speech pathologist who had diverse experiences with clients of different ages, including clinical experience within adult and pediatric tertiary multidisciplinary pain centers (authors NEA, DW, and MS). In CSIRO’s previous chatbot development [24,26], the personality of the person entering the data was evident in the chatbot through subtle language cues; users disengaged quickly unless the programmer who wrote the script was both knowledgeable and personable. Having clinicians write responses for their respected client base provided consistency and portrayed the chatbot as a warm conversation partner with a unique personality.

The initial chatbot education topics were mapped by the interdisciplinary research team with clinical backgrounds in pain management, as illustrated in Figure 2. Commonly asked questions on these topics were then identified, and the subsequent topics, questions, and conversation flows were discussed. First, an adult version of the app content was created. Where possible, existing resources for information were referred to in Dolores’ responses, including web pages, YouTube videos, mobile phone apps, and web-based education modules. This content was then altered to ensure age appropriateness for older and younger adolescents by a pediatric speech pathologist researcher (author TR) and pediatric pain clinicians (authors DW and MS). These alterations included lowering the readability level of the text and referring to more age-appropriate resources. An example of the response to a specific question posed by users of different ages (ie, adults, older adolescents, and younger adolescents) can be seen in Table 1. Implementation proceeded according to the agreed chatbot specifications, and responses were compiled into the brain by an experienced computer scientist (author DI), with assistance from an intern (author PV).

For the purpose of pilot testing, Dolores was built to be used as a stand-alone iOS or iPadOS app. The user interface allowed the clinician to select the language level (ie, adult, older adolescent, or younger adolescent) to match the users’ language proficiency. Dolores generated responses using a combination of speech and written outputs with a “chat bubble” produced as typically seen in a conventional web-based chatting program.

The speech output was set to the VoiceOver setting selected on the device. The default setting of Siri Female (Australia) with a pitch change of 50% and speaking rate of 50% was used for pilot testing. An audio recording of the speech output can be found in Multimedia Appendix 1. The user could interact with Dolores through text, speech, and optional interactive buttons. User speech recognition used Google Cloud speech to text, as it has the highest accuracy of available programs and provides a nonlogging option to ensure privacy [28].

On the basis of previous research emphasizing the importance of being able to tell one’s story to enhance the experience of pain education [11], 2 interactive experiences were built. A structured pain history interview was first administered by Dolores, which allowed users to begin developing rapport with Dolores and reflect on their own experiences. Key assessment constructs for the pain history interview were drawn from existing assessment proformas used by local pain centers and the electronic Persistent Pain Outcomes Collaboration (ePPOC) referral questionnaires, which are widely used across Australia and New Zealand [29,30]. The assessment constructs included pain location, pain duration, sensory descriptions of pain, affective descriptors of pain, initial trigger or beginning of the pain experience, perceived impact of pain on daily life, current understanding of pain, and perceptions of treatment approaches. Users could select a drawing widget to answer questions related to pain location, descriptors of pain, and how their pain began. A pain diagram could also be used to mark the location of pain.

The second component of the chatbot interaction is the educational session. Pain education was initiated by Dolores asking the user about their learning preferences. Users were able to browse topics, select a random topic, take a quiz, or ask Dolores questions. If a user wished to continue learning about a topic after the chatbot interaction, the information provided by the Dolores could be saved and then exported to a PDF format for future reference. The chatbot was tested multiple times by the research team before pilot testing with people with chronic pain to verify the appropriateness of the content and conversation flow. Various changes were made to the app layout and content before pilot testing with people with chronic pain (eg, shortening the length of responses and modification of interface labels). Screenshots of the Dolores interface are shown in Figure 3.

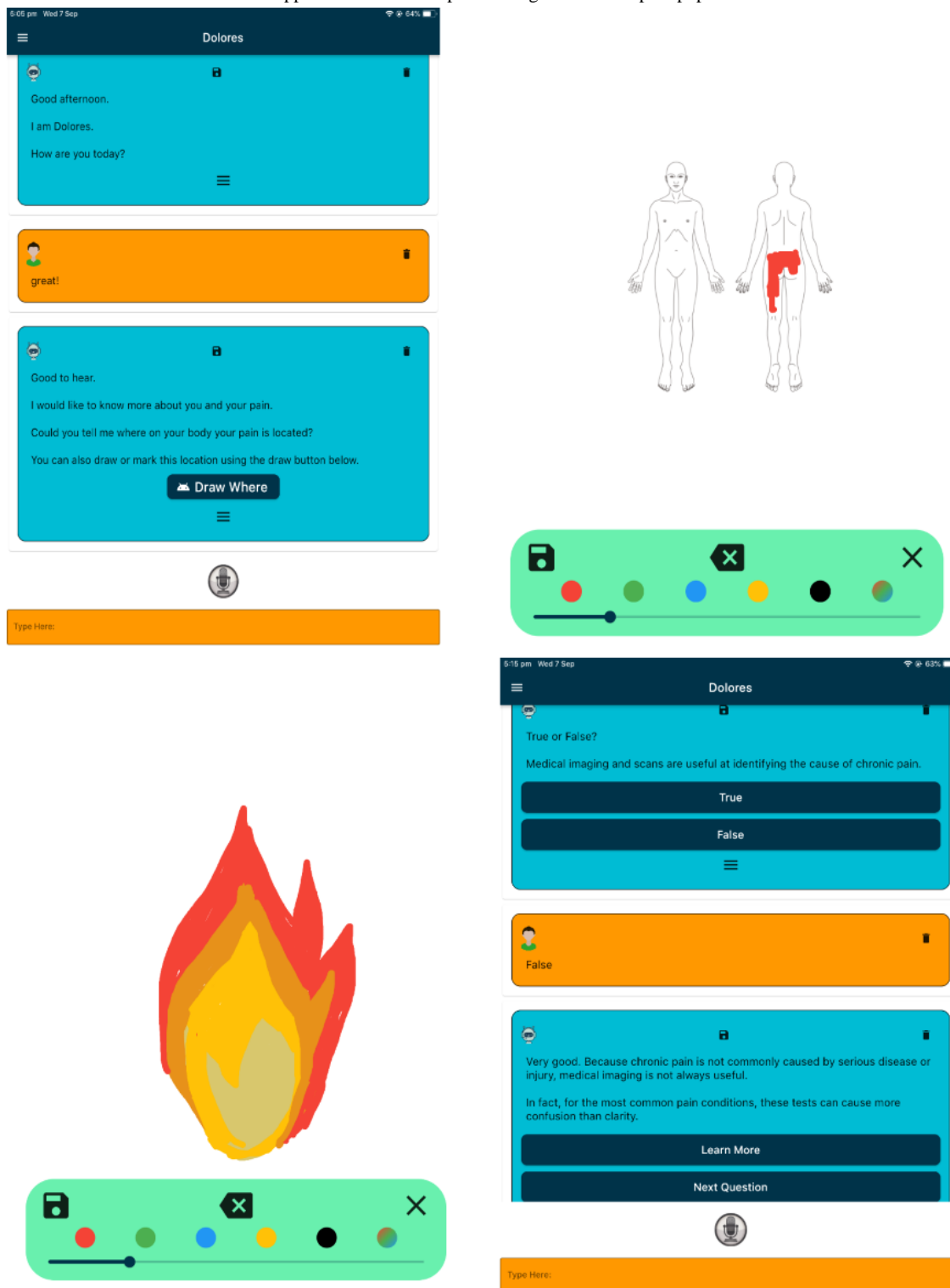
Figure 2. Initial Dolores chatbot education topics that were mapped by the interdisciplinary research team for pilot testing in a chronic pain population.



Table 1. Example of the Dolores chatbot age-appropriate responses to a user’s question that was developed for pilot testing in a chronic pain population.

Response to the question posed by a user: what is mindfulness?	
Response to adults	Mindfulness is about being present in the moment, accepting the thoughts and feelings you’re having without worrying about them, allowing you to achieve a sense of peace. Essentially, mindfulness therapies teach you to be aware of the pain you are feeling and the thoughts you may be having about that pain, without attaching any negative connotations to them. More information about mindfulness can be found at this website [31]. If you want to trial mindfulness, Smiling Mind is a free mindfulness and relaxation app [32]. For more details, visit the Smiling Mind website.
Response to older adolescents	Mindfulness is a skill you can learn to help notice thoughts and feelings in order to feel comfortable and safe. There are many ways to learn mindfulness strategies and your therapist might help you learn how to use this skill as a part of your pain treatment. Talk to your pain team about mindfulness if you would like to know more.
Response to younger adolescents	You can learn how to become more aware of your thoughts and feelings. This is a skill called mindfulness. It can help you to feel relaxed and safe. There are a lot of different things you can do to help you notice what your thoughts and feelings are right now. You might learn how to be more mindful as part of your pain treatment. You can talk to your pain team if you want to know more about mindfulness.

Figure 3. Screenshots of the Dolores chatbot app interface used for pilot testing in a chronic pain population.



Ethical Considerations

The Children’s Health Queensland Hospital and Health Service Human Research Ethics Committee (HREC/21/QCHQ/73157), the CSIRO Health and Medical Human Research Ethics Committee (2021_049_RR), and the University of Queensland Human Research Ethics Committee (2021/HE001056) approved the protocol for pilot testing as detailed further. Written informed consent was obtained before participating in the study.

For younger participants (aged <18 years), consent was obtained from both the participants and their parents or guardians. All data were deidentified at the time of data collection; participants at each site were allocated a unique research ID number to identify all study materials. All the data were stored on a secure server at the University of Queensland (UQ Research Data Manager). At the end of the study, participants received an Aus \$25 (US \$18.75) voucher for use in popular retail stores in Australia as a gesture of appreciation. To avoid coercion,

participants were not informed of the voucher until the end of the study.

Pilot Testing With People With Chronic Pain

Recruitment and Procedure

Participants were recruited from 1 of 2 pain centers located in large tertiary hospitals in Australia. The inclusion criteria were as follows: the participant must (1) have persistent noncancer pain for at least 3 months; (2) be aged ≥ 10 years; and (3) possess adequate expressive and receptive language skills to complete the required tasks in English, as determined by clinicians in the treatment team. Participants with coexisting medical conditions were not excluded from this study. As little was known about the variables of interest in this study, the target sample size was based on the rules of thumb for regression (number of predictor variables+50) [33]. A target sample size of 60 participants was used to account for any missing data.

Participants meeting the inclusion criteria were identified by 1 of the 2 physiotherapy research students (authors LB and EH) on presentation to their outpatient appointment. All participants were provided with written and verbal information regarding the study. The research students emphasized that participation was voluntary and that decisions about participation would not affect the patient's relationship with the clinical staff or their ongoing care at the pain center. Consenting participants were transferred to a private treatment space and completed a short paper-based demographic questionnaire. Participants were then given a fifth-generation iPad with the Dolores app installed. One of the research students provided a 2-minute orientation to the app and provided participants with a cheat sheet (Multimedia Appendix 2). Participants then spent 20 to 30 minutes completing the 2 interactive chatbot activities (ie, the pain history interview and education session). The research student remained in the room during the chatbot interaction to provide technical assistance or answer any questions. However, minimal assistance was required during the pilot testing. On completion of the chatbot experience, participants completed a paper-based custom-made evaluation questionnaire. Adolescent participants were allowed to have their parents or guardians present to assist them in completing all tasks.

Measures

Demographic Questionnaire

Information regarding participant gender, age, country of birth, languages spoken, language spoken at home, education level, coexisting medical diagnoses affecting communication and language, pain onset, pain duration, and pain location were gathered. Confidence in using mobile phone apps was measured on a 5-point Likert scale (0=not at all confident; 4=very confident). An estimation of the amount of time spent using a smartphone in the last 7 days was also recorded.

Custom-Made Evaluation Questionnaire

A custom-made evaluation questionnaire that explored chatbot acceptability was created to ensure that the questions were suitable and easy to understand across all the age groups. Eight acceptability constructs were chosen based on scoping reviews focused on either the technical evaluation of health chatbots or

the measurement of mobile health acceptability [34,35]. The 8 constructs examined were engagement, perceived value, usability, accuracy, responsiveness, esthetics, adoption intention, and overall quality. All constructs were rated using a 5-point Likert agreement scale with pictorial representations (ie, emoji faces). Furthermore, 4 open-ended questions were included to allow feedback on the most- and least-liked features and suggestions for future updates. The evaluation questionnaire is presented in Multimedia Appendix 3.

Statistical Analysis

All statistical analyses were conducted using SPSS for Windows (version 28.0.1; IBM Corp). Descriptive statistics were first generated to describe the sample and examine feedback pertaining to the Likert-scale acceptability ratings from the evaluation questionnaire. Frequencies were generated for the questionnaire scale responses. A series of cumulative odds ordinal logistic regression models with proportional odds was then generated to determine the effect of age group on acceptability ratings.

Responses to the open-ended questions of the evaluation questionnaire were analyzed using applied thematic analysis with an exploratory approach [36]. Text segmentation was not used, and responses to all the questions were considered in their entirety when generating overarching themes. The text was first analyzed by a researcher with a clinical background in pain management and experience in qualitative analysis (author NA) [37-39] to develop codes and overarching themes. Both codes and themes were reviewed by a second researcher with a background in computer science and experience in qualitative analysis (author DI) [26]. The number of participants whose comments were coded with a particular code (ie, frequency count) was then generated.

The high-frequency codes across the entire sample were examined further. A threshold of a frequency count >10 was used to identify high-frequency codes, and a comparative analysis was undertaken on the high-frequency codes that were distinct from the constructs examined using acceptability Likert scales. Binomial logistic regression models were generated to ascertain the effects of age group on the likelihood of these unique high-frequency codes appearing in the text of open-ended question feedback. An additional binomial logistic regression model was used to establish the effect of age group on providing any comments in the evaluation questionnaire. The following cutoff points were used to interpret the effect size for odds ratios: 1.68 or 0.60 (small), 3.47 or 0.29 (medium), and 6.71 or 0.15 (large) [40].

Results

A total of 60 individuals with chronic pain were recruited to participate in this study: 20 (33%) adolescent participants (aged 10-18 years), 20 (33%) young adults (aged 19-35 years), and 20 (33%) adults (aged >35 years) with chronic pain.

Demographic Data

The demographic information pertaining to each age group and the sample as a whole is presented in Table 2. The mean ages of the adolescent, young adult, and adult participants were 15.05

(SD 1.91; range 10-18), 28.79 (SD 4.6; range 21-35) and 54.6 (SD 12.0; range 38-76) years, respectively. More female (14/20, 70%) than male (6/20, 30%) adolescent participants were recruited for the study. An equal number of male and female individuals participated in the other age cohorts. Most participants (54/59, 92%) were born in Australia and identified English as their primary language. The adult group had the highest prevalence of participants (4/20, 20%) who spoke multiple languages. Self-reported coexisting medical conditions known to affect language and communication (eg, dyslexia)

were low across the age groups. Overall, the most common pain locations were the lower back (30/59, 51%), upper limbs (25/59, 42%), and neck (25/59, 39%), and medical conditions were the most common cause of pain onset identified (18/59, 31%). The adolescent cohort reported being the most confident in using apps, with 75% (15/20) of participants selecting “very confident.” The time spent using a smartphone was relatively similar across the age groups, with a mean of 25.77 (SD 18.52) hours per week recorded for the sample.

Table 2. Demographic information for the chronic pain study cohorts who participated in pilot testing of the Dolores chatbot app.

Demographic variable	Total (n=59)	Adults (n=20)	Young adult (n=19)	Adolescents (n=20)
Gender, n (%)				
Male	26 (44)	10 (50)	10 (53)	6 (30)
Female	33 (56)	10 (50)	9 (47)	14 (70)
Age (years), mean (SD; range)	32.88 (18.24; 10-76)	54.6 (12.00; 38-76)	28.79 (4.59; 21-35)	15.05 (1.91; 10-18)
Country of birth, n (%)				
Australia	54 (92)	16 (80)	19 (100)	19 (95)
Egypt	1 (2)	1 (5)	0 (0)	0 (0)
England	1 (2)	1 (5)	0 (0)	0 (0)
New Zealand	1 (2)	1 (5)	0 (0)	0 (0)
The Netherlands	1 (2)	1 (5)	0 (0)	0 (0)
Germany	1 (2)	0 (0)	0 (0)	1 (5)
Languages spoken, n (%)				
English only	54 (92)	16 (80)	19 (100)	19 (95)
Multiple (including English)	5 (9)	4 (20)	0 (0)	1 (5)
Main language, n (%)				
English	59 (98)	19 (95)	19 (100)	20 (100)
Arabic	1 (2)	1 (5)	0 (0)	0 (0)
Education level, n (%)				
University	11 (19)	6 (30)	5 (46)	0 (0)
Diploma	6 (10)	4 (20)	2 (11)	0 (0)
Certificate	13 (22)	5 (25)	6 (32)	2 (10)
Secondary education	20 (34)	5 (25)	6 (32)	9 (45)
Primary education	9 (15)	0 (0)	0 (0)	9 (45)
Coexisting medical conditions, n (%)				
Dyslexia	3 (5)	1 (5)	0 (0)	2 (10)
Vision impairment	2 (3)	2 (10)	0 (0)	1 (5)
Behavior or attention disorder	4 (7)	0 (0)	2 (11)	2 (10)
Autism spectrum disorder	2 (3)	0 (0)	1 (5)	1 (5)
Pain onset, n (%)				
Injury at home	3 (5)	1 (5)	1 (5)	1 (5)
Injury at work or school	9 (15)	3 (15)	4 (21)	2 (10)
Injury in another setting	8 (14)	3 (15)	4 (21)	1 (5)
Road traffic crash	5 (9)	3 (15)	1 (5)	1 (5)
Medical condition	18 (31)	8 (40)	5 (26)	5 (25)
After surgery	5 (9)	1 (5)	2 (11)	2 (10)
No obvious cause	11 (19)	1 (5)	2 (11)	8 (40)
Pain duration (years), mean (SD; range)	8.09 (8.09; 0.67-40)	15.00 (12.26; 0.67-40)	6.05 (5.51; 1-20.5)	3.12 (3.09; 0.67-14.5)
Pain location, n (%)				
Head or face	12 (20)	5 (25)	2 (11)	5 (25)
Lower back	30 (51)	12 (60)	10 (53)	8 (40)
Upper limb	25 (42)	9 (45)	6 (32)	10 (50)
Abdomen or groin	15 (25)	4 (20)	4 (21)	7 (35)

Demographic variable	Total (n=59)	Adults (n=20)	Young adult (n=19)	Adolescents (n=20)
Hips	17 (29)	8 (40)	5 (26)	4 (20)
Knees	15 (25)	7 (35)	3 (16)	5 (25)
Neck	23 (39)	10 (50)	5 (26)	8 (40)
Upper back	16 (27)	3 (15)	6 (32)	7 (35)
Chest	7 (12)	2 (10)	2 (11)	3 (15)
Buttocks	4 (7)	3 (15)	1 (5)	0 (0)
Thighs	7 (12)	3 (15)	2 (11)	2 (10)
Calves, ankles, or feet	12 (20)	5 (25)	3 (16)	4 (20)
Total body	2 (3)	1 (5)	1 (5)	0 (0)
Confidence using apps, n (%)				
Not at all confident	0 (0)	0 (0)	0 (0)	0 (0)
Slightly confident	1 (2)	1 (5)	0 (0)	0 (0)
Moderately confident	9 (15)	8 (40)	1 (5)	0 (0)
Confident	20 (34)	8 (40)	36.8 (7)	5 (25)
Very confident	29 (49)	3 (15)	11 (58)	15 (75)
Hours using smartphone, mean (SD; range)	25.77 (18.52; 0-70) ^a	20.58 (15.13; 2-56) ^b	28.68 (18.88; 3-70)	27.93 (20.84; 0-70)

^aOne adult nonrespondent, thus n=58.

^bOne adult nonrespondent, thus n=19.

Likert Scale Acceptability Ratings

The acceptability ratings for the entire sample are presented in Table 3. Overall, the acceptability was high. A very small number of participants (3/57, 5%) indicated that they either disagreed or strongly disagreed with one of the statements. The usability and responsiveness ratings were particularly high, with most participants rating these statements as either agree or strongly agree (54/57, 95% of the sample for usability and 52/57, 91% of the sample for responsiveness). Perceived value was

the lowest-rated statement, with 32% (18/57) of the sample recording a neutral response.

The ordinal logistic regression models examining the effect of age group on acceptability ratings are presented in Table 4. The effect of age group on all acceptability ratings was small and not statistically significant. The assumption of proportional odds was met for all models, as assessed by nonsignificant full likelihood ratio tests comparing the fit of the proportional odds model to a model with varying location parameters.

Table 3. Acceptability ratings across the whole sample (n=57) based on pilot testing data of the Dolores chatbot app in a chronic pain population.

Acceptability construct	Acceptability Rating					
	Strongly Agree, n (%)	Agree, n (%)	Somewhat Agree, n (%)	Neutral, n (%)	Disagree, n (%)	Strongly Disagree, n (%)
Engagement: I enjoyed talking to Dolores	20 (35)	29 (51)	0 (0)	7 (12)	0 (0)	1 (2)
Usability: It was easy to talk to Dolores	22 (39)	32 (56)	0 (0)	3 (5)	0 (0)	0 (0)
Responsiveness: Dolores was fast enough when responding back to me	30 (53)	22 (39)	0 (0)	5 (9)	0 (0)	0 (0)
Esthetics: I liked the design of the Dolores app including graphics and layout	21 (37)	25 (44)	2 (4)	8 (14)	0 (0)	1 (2)
Perceived value: Dolores helped me to understand more about my pain or my pain treatments	9 (16)	26 (46)	0 (0)	20 (35)	2 (4)	0 (0)
Accuracy: Dolores understood what I was asking or saying	24 (42)	22 (39)	0 (0)	9 (16)	2 (4)	0 (0)
Adoption intention: I would talk to Dolores again	19 (33)	29 (51)	0 (0)	8 (14)	0 (0)	1 (2)
Overall quality: I would recommend Dolores to others	25 (45)	24 (43)	0 (0)	6 (11)	0 (0)	1 (2)

Table 4. Logistic Ordinal Regression models investigating acceptability ratings across age groups (n=57) based on pilot testing data of the Dolores chatbot app in a chronic pain population.

Statistic	Engagement	Perceived value	Usability	Accuracy	Responsiveness	Adoption intention	Esthetics	Overall quality
Pediatric								
OR ^a (95% CI)	2.33 (0.69-7.94)	2.56 (0.76-8.56)	2.68 (0.73-9.78)	1.72 (0.53-5.57)	0.45 (0.27-3.10)	2.67 (0.76-9.19)	1.98 (0.61-6.47)	2.34 (0.69-8.07)
P value	.17	.13	.14	.37	.90	.12	.26	.17
Young adult								
OR (95% CI)	1.82 (0.53-6.23)	2.86 (0.84-9.79)	1.15 (0.33-4.02)	1.23 (0.37-4.07)	0.76 (0.22-2.65)	0.88 (0.26-2.98)	0.60 (0.18-2.02)	1.45 (0.42-4.95)
P value	.34	.09	.83	.74	.67	.84	.41	.55
Adult^b								
OR (95% CI)	1.00 (N/A ^c)	1.00 (N/A)	1.00 (N/A)	1.00 (N/A)	1.00 (N/A)	1.00 (N/A)	1.00 (N/A)	1.00 (N/A)
Chi-square (df)	2.0 (4)	3.6 (4)	2.6 (2)	0.8 (4)	17.1 (2)	3.7 (4)	3.6 (2)	1.9 (4)
P value	.37	.17	.27	.66	.91	.16	.17	.39

^aOR: odds ratio. Higher odds ratios indicate that the age group is more likely to provide negative ratings.

^bReference category.

^cN/A: not applicable.

Qualitative Feedback

All codes and overarching themes generated from the applied thematic analysis of responses to open-ended questions from the evaluation questionnaire are displayed in Table 5. Most participants (52/57, 91%) provided comments about their experiences by responding to the open-ended questions. The average number of words used to respond to the open-ended questions across the sample was 10.67 (SD 10.22; range 0-56). An analysis of all responses to all questions revealed three overarching themes related to participants' experiences interacting with Dolores: (1) frustrations, (2) refinements, and (3) benefits.

Most codes relating to the theme *frustrations* concerned Dolores' speech output; 9 participants expressed that Dolores' voice was unpleasant or too robotic, 7 participants commented that the speech rate was too slow, and 2 participants highlighted that the pronunciation of certain words was incorrect: "She was a slow talker" (Adolescent participant 12) and "She says words funny" (Adolescent participant 8).

For the theme *refinements*, participant codes encompassed suggestions on how the Dolores app could be improved; 8 participants advocated for adjustable voice options and speech rate settings for Dolores: "a setting to change the speed of talking" (Adolescent participant 11). A further 2 participants recommended an optional mute function for Dolores' interaction with the users: "have talking be an option" (Adolescent participant 16). Other recommendations that were endorsed by >1 participant included more in-depth or specific information on certain topics, more quiz questions, and the ability to use the app at home.

Under the theme *benefits*, a variety of codes reflected feedback about positive experiences using the Dolores app; 48 participants commented on their positive experiences, whereas 19 participants indicated that the best part of the experience was that the app was easy to use: "easy to use and interact with" (Young Adult participant 2). The efficiency of the app, the layout, and the design were mentioned by 4 participants. The ease with which participants were able to disclose information to Dolores was also highlighted by 4 participants. In addition, 12 participants reported that the information provided by Dolores was useful: "information was brief but helpful" (Young Adult participant 2), whereas 7 participants commented favorably on the multiple options available for user interaction: "I liked the variety of ways to communicate" (Adolescent participant 15).

The different age groups were equally likely to provide feedback by responding to the open-ended questions in the evaluation questionnaire. A binomial logistic regression model examining the effect of age group on providing comments was not statistically significant, and the effect sizes were small ($\chi^2_2=0.5$; $P=.78$; Nagelkerke $R^2=0.02$). Dolores' speech (code A) was considered both a high-frequency code and distinct from the constructs examined using the acceptability Likert scales. The comparative analysis examining the effect of age group on providing negative comments about Dolores' speech was statistically significant ($\chi^2_2=11.7$; $P=.003$). The model explained 27.1% (Nagelkerke R^2) of the variance. Adults and young adults were less likely to comment on Dolores' speech compared with adolescent participants. Using the adolescent cohort as the reference group, a medium effect size was observed for adults (odds 77% lower), and a large effect size was observed for young adults (odds 95% lower).

Table 5. Coding and themes from the evaluation questionnaire (n=57) based on pilot testing of the Dolores chatbot app in a chronic pain population.

	Frequency counts ^a , n (%)
Theme A: frustrations	19 (33)
Code A: Dolores' speech	18 (32)
Code A1: unpleasant voice	9 (16)
Code A2: speech rate too slow	7 (12)
Code A3: incorrect pronunciation of words	2 (4)
Code B: unfamiliar	2 (4)
Code C: childlike	1 (2)
Code D: having to type on a touch screen	1 (2)
Code E: too assertive	1 (2)
Theme B: refinements	24 (42)
Code F: adjustable voice or speech settings	8 (14)
Code G: more specific or in-depth information	3 (5)
Code H: more questions in quiz	2 (4)
Code I: home access or being able to use for longer	2 (4)
Code J: able to turn off speech option for Dolores	2 (4)
Code K: add a skip button	1 (2)
Code L: color codes for pain severity on body map	1 (2)
Code M: avatar for Dolores	1 (2)
Code N: more questions in pain history	1 (2)
Code O: more links to existing resources	1 (2)
Code P: edit button for typed entries	1 (2)
Code Q: improved speech to text	1 (2)
Code R: Bluetooth keyboard	1 (2)
Code S: easier to erase when drawing	1 (2)
Code T: monthly tracking of progress	1 (2)
Code U: additional games related to topics	1 (2)
Theme C: benefits	48 (84)
Code V: easy to use	19 (33)
Code W: information useful	12 (21)
Code W1: informative	10 (18)
Code W2: helpful information	1 (2)
Code W3: information had not been provided by health professionals	1 (2)
Code X: options for user to interact, including drawing	7 (12)
Code X1: interactive	2 (4)
Code X2: multiple ways to respond	3 (5)
Code X3: being able to draw responses	2 (4)
Code X4: being able to speak to app	1 (2)
Code Y: quick or efficient	4 (7)
Code Z: easy to open up about issues or talk to	4 (7)
Code AA: good layout or design	4 (7)
Code AB: good tool or generalized app	3 (5)
Code AC: mobile or easy to access	1 (2)

	Frequency counts ^a , n (%)
Code AD: asked questions that may not think of	1 (2)
Code AE: real-time response or feedback	1 (2)
Code AF: interesting quiz	1 (2)
Code AG: lets user make mistakes	1 (2)
Code AH: speech to text accuracy	1 (2)

^aNumber of participants whose comments were coded with a particular code or theme.

Discussion

Principal Findings

The utility of chatbots in the health care sector is an emerging field of research. This study details the development and initial pilot testing of a multimodality pain history assessment and pain education chatbot (Dolores) designed to be used across different age groups from early adolescence to adulthood. We hypothesized that the acceptability ratings and feedback would be comparable across different age groups following initial pilot testing.

The overall acceptability ratings were high across the 3 age groups examined (ie, adolescents, young adults, and adults), and the effect of age group on all acceptability ratings was small and not statistically significant. The positive feedback provided to the open-ended questions also outweighed the negative or constructive feedback. The participants provided positive feedback on the multiple options available for interaction and commented that the app was easy to use. Dolores was designed to provide age-appropriate responses in terms of both content and readability level using 3 settings (ie, for younger adolescents, older adolescents, and adults). Dolores was also programmed to accommodate different user communication needs (eg, visual, speech, and text) and learning preferences (eg, taking a quiz, browsing topics, and asking questions). The current settings and options appeared appropriate, as only one participant provided negative feedback regarding the app not being targeted to the participants' age group (ie, an adult stating that the app was a bit childlike). This is in contrast to a previous study that has revealed that adolescents often feel that health apps are designed for adults, impacting adoption intention [41]. The results of this study suggest that multimodality pain education chatbots are suitable for use across the different age groups examined.

Although most feedback was positive, 30% (18/57) of participants made specific negative comments about Dolores' speech, which was set to Siri Female (Australia) with a pitch change of 50% and a speaking rate of 50% (Multimedia Appendix 1). Participants reported that the speech rate was too slow, the voice output was unpleasant, or Dolores pronounced words incorrectly. Adults and young adults were statistically less likely to comment on Dolores' speech compared with adolescent participants. However, the reason for this finding remains unclear. As all age groups were equally likely to provide comments in response to the open-ended questions, the results could not be adequately explained by response bias.

One possible explanation for age-related differences in opinions pertaining to Dolores' speech may be related to human-robot age congruity. Previous experimental research within the travel industry has demonstrated that human-robot gender congruity elicits more positive affect whereby participants are more likely to rate their level of comfort in the service encounter higher if the robot is perceived to be of the same gender as the participant [42]. Similarly, in the retail industry, research has demonstrated that matching consumer personality with congruent chatbot personality has a positive impact on both consumer engagement with chatbots and purchasing outcomes [43]. As the voice selected for pilot testing resembled an adult's voice, it may be that adolescent participants were more likely to comment on Dolores' speech due to age-related incongruities between the participant's own voice and Dolores'.

Alternatively, the desired level of anthropomorphism may explain the association between age and negative comments about Dolores' speech. A narrative review of literature comparing embodied with disembodied chatbots, used in a variety of research fields, yielded contradictory results regarding the effect of human likeness on user acceptability [44]. Current research suggests that extreme anthropomorphic features may lead to cognitive dissonance, whereas a chatbot with no human identity could affect perceptions of trust and the quality of the interaction [16]. However, research examining preferences within pediatric and adolescent cohorts is lacking. It is possible that there is a preference for more fictional personas, such as animals or mythical creatures, within this age cohort. The Kids and Family Reading Report [45], which sampled 2758 parents and their children, revealed that children and adolescents aged 6 to 17 years are equally likely to learn life lessons from fictional and nonfictional characters [45]. The types of characters children and adolescents want in books reflect their own desired personal attributes such as resilience, intelligence, bravery, and strength [45]. No known research has explored the impact of using different chatbot personas with adolescents or children in the health field.

It is unclear how negative perceptions of a chatbot's speech might impact other acceptability constructs; there were no age-related differences in the acceptability ratings in this study for the construct engagement, perceived value, usability, accuracy, responsiveness, adoption intention, esthetics, and overall quality. It is possible that negative perceptions of chatbots' speech and persona might impact perceived trustworthiness and long-term engagement, which were not measured. In clinical settings, improved treatment outcomes have been linked to the strength of the therapeutic alliance between therapists and people seeking treatment for chronic

pain [46], and perceived trustworthiness has been shown to be important in therapeutic encounters [47]. Participants recommended several minor adaptations to the Dolores app, including mute options and adjustable voice and speech settings. Although these adaptations may improve acceptability, further research is needed to examine chatbot personality characteristics that can optimize trustworthiness and long-term engagement in adolescents.

Incorporating technology into health care for young people is becoming increasingly recognized as an avenue for enhancing care [48,49]. Long-term engagement with chatbots has the potential to increase the ways young people can interact with their health care and allow young people to exercise more control over decisions related to their health [50,51]. Prior research has found that young people aged <17 years exhibit a mean of only 15% of all utterances during a specialist medical appointment and that active participation is increased when young people are engaged directly through a range of strategies that allow them to speak about their daily life, make choices, and indicate their preferences [52]. Chatbots may be highly valuable tools for increasing engagement, communication, understanding, and clinical outcomes in adolescents. Therefore, further research on long-term chatbot use as an adjunct to existing evidence-based interventions and services is warranted.

Limitations

The results of this study should be interpreted with consideration of the following limitations. The sample size was small, and only 2 clinical sites were used to recruit participants, limiting the generalizability of the findings. A custom-made evaluation questionnaire was created to ensure that the measurement was suitable and easy to understand across all age groups. However, this may have introduced measurement errors that affected the confidence of the findings. A researcher was present during the

pilot testing to assist with any technical issues. This may have affected chatbot engagement and acceptability due to the Hawthorne effect. Participants were, however, explicitly told that the researchers were only present to provide technical support, and they were positioned out of direct sight to decrease the influence of their presence on the participants' behavior. Broad categorizations were used for age grouping to facilitate recruitment; however, definitions for "adolescent" and "young adult" do vary. Children (ie, aged <10 years) and an older adult group (eg, exclusively those aged >65 years) were not recruited due to perceived difficulties in recruiting these cohorts at the clinical sites participating in the study. All the participants in this study indicated that they were at least slightly confident when using smartphone apps. Younger children or older adults may not have had any exposure to apps, with the potential to reduce acceptability in those age groups. Investigating the acceptability of pain education chatbots with older adults, children aged <10 years, and populations with lower health and digital literacy skills is an avenue for future research.

Conclusions

This study detailed the development and pilot testing of a multimodality chatbot within the pain field. The study provided support for the acceptability of pain education chatbots for conducting a pain history assessment and delivering education across age groups spanning from early adolescence to adulthood. Adolescent participants were more likely to provide specific negative comments on the chatbot's speech compared with adults and younger adults. Although chatbot acceptability may be improved by enabling the self-selection of voice and speech settings, more research is needed to understand how the embodiment of chatbots can be optimized, with consideration of audiovisual preferences relating to age and gender, to facilitate long-term engagement as a complement to standard care.

Acknowledgments

PV was affiliated with CSIRO AEHRC at the time of the study, and is currently affiliated with the Walter and Eliza Hall Institute of Medical Research.

Data Availability

The data sets generated and analyzed during this study are available from the corresponding author on reasonable request.

Authors' Contributions

All authors attest that they met the International Committee of Medical Journal Editors criteria for authorship. NEA, DI, DW, TR, MS, JS, and AC made substantial contributions to the conception and design of this study. NEA, DI, PV, DW, TR, MW, JS, and AC contributed to the development of Dolores (brain or app design) and study materials. LB and EH substantially contributed to the data collection under the supervision of NEA, DI, DW, and AC. NEA and DI analyzed and interpreted the data. NEA and DI drafted the manuscript. All authors discussed the results and revised and approved the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Audio sample of Dolores' speech from pilot testing the Dolores chatbot app in a chronic pain population.

[[MP4 File \(MP4 Video\)](#), 349 KB-Multimedia Appendix 1]

Multimedia Appendix 2

Orientation to the Dolores app: cheat sheet that was used for pilot testing in a chronic pain population.

[\[PDF File \(Adobe PDF File\), 295 KB-Multimedia Appendix 2\]](#)

Multimedia Appendix 3

A custom-made evaluation questionnaire assessing acceptability constructs relevant to chatbots that was used for the pilot testing of the Dolores chatbot app in a chronic pain population.

[\[PDF File \(Adobe PDF File\), 59 KB-Multimedia Appendix 3\]](#)

References

1. Treede RD, Rief W, Barke A, Aziz Q, Bennett MI, Benoliel R, et al. Chronic pain as a symptom or a disease: the IASP Classification of Chronic Pain for the International Classification of Diseases (ICD-11). *Pain* 2019 Jan;160(1):19-27 [doi: [10.1097/j.pain.0000000000001384](https://doi.org/10.1097/j.pain.0000000000001384)] [Medline: [30586067](https://pubmed.ncbi.nlm.nih.gov/30586067/)]
2. Cohen SP, Vase L, Hooten WM. Chronic pain: an update on burden, best practices, and new advances. *Lancet* 2021 May 29;397(10289):2082-2097 [doi: [10.1016/S0140-6736\(21\)00393-7](https://doi.org/10.1016/S0140-6736(21)00393-7)] [Medline: [34062143](https://pubmed.ncbi.nlm.nih.gov/34062143/)]
3. GBD 2016 Disease and Injury Incidence and Prevalence Collaborators. Global, regional, and national incidence, prevalence, and years lived with disability for 328 diseases and injuries for 195 countries, 1990-2016: a systematic analysis for the Global Burden of Disease Study 2016. *Lancet* 2017 Sep 16;390(10100):1211-1259 [FREE Full text] [doi: [10.1016/S0140-6736\(17\)32154-2](https://doi.org/10.1016/S0140-6736(17)32154-2)] [Medline: [28919117](https://pubmed.ncbi.nlm.nih.gov/28919117/)]
4. St Sauver JL, Warner DO, Yawn BP, Jacobson DJ, McGree ME, Pankratz JJ, et al. Why patients visit their doctors: assessing the most prevalent conditions in a defined American population. *Mayo Clin Proc* 2013 Jan;88(1):56-67 [FREE Full text] [doi: [10.1016/j.mayocp.2012.08.020](https://doi.org/10.1016/j.mayocp.2012.08.020)] [Medline: [23274019](https://pubmed.ncbi.nlm.nih.gov/23274019/)]
5. King S, Chambers CT, Huguette A, MacNevin RC, McGrath PJ, Parker L, et al. The epidemiology of chronic pain in children and adolescents revisited: a systematic review. *Pain* 2011 Dec;152(12):2729-2738 [doi: [10.1016/j.pain.2011.07.016](https://doi.org/10.1016/j.pain.2011.07.016)] [Medline: [22078064](https://pubmed.ncbi.nlm.nih.gov/22078064/)]
6. Sharpe L, Jones E, Ashton-James CE, Nicholas MK, Refshauge K. Necessary components of psychological treatment in pain management programs: a Delphi study. *Eur J Pain* 2020 Jul 30;24(6):1160-1168 [doi: [10.1002/ejp.1561](https://doi.org/10.1002/ejp.1561)] [Medline: [32187442](https://pubmed.ncbi.nlm.nih.gov/32187442/)]
7. Siddall B, Ram A, Jones MD, Booth J, Perriman D, Summers SJ. Short-term impact of combining pain neuroscience education with exercise for chronic musculoskeletal pain: a systematic review and meta-analysis. *Pain* 2022 Jan 01;163(1):e20-e30 [doi: [10.1097/j.pain.0000000000002308](https://doi.org/10.1097/j.pain.0000000000002308)] [Medline: [33863860](https://pubmed.ncbi.nlm.nih.gov/33863860/)]
8. Robins H, Perron V, Heathcote LC, Simons LE. Pain neuroscience education: state of the art and application in pediatrics. *Children (Basel)* 2016 Dec 21;3(4):43 [FREE Full text] [doi: [10.3390/children3040043](https://doi.org/10.3390/children3040043)] [Medline: [28009822](https://pubmed.ncbi.nlm.nih.gov/28009822/)]
9. Moseley GL, Butler DS. Fifteen years of explaining pain: the past, present, and future. *J Pain* 2015 Sep;16(9):807-813 [doi: [10.1016/j.jpain.2015.05.005](https://doi.org/10.1016/j.jpain.2015.05.005)] [Medline: [26051220](https://pubmed.ncbi.nlm.nih.gov/26051220/)]
10. Koechlin H, Locher C, Prchal A. Talking to children and families about chronic pain: the importance of pain education—an introduction for pediatricians and other health care providers. *Children (Basel)* 2020 Oct 12;7(10):179 [FREE Full text] [doi: [10.3390/children7100179](https://doi.org/10.3390/children7100179)] [Medline: [33053802](https://pubmed.ncbi.nlm.nih.gov/33053802/)]
11. Watson JA, Ryan CG, Cooper L, Ellington D, Whittle R, Lavender M, et al. Pain neuroscience education for adults with chronic musculoskeletal pain: a mixed-methods systematic review and meta-analysis. *J Pain* 2019 Oct;20(10):1140.e1-1140.e22 [doi: [10.1016/j.jpain.2019.02.011](https://doi.org/10.1016/j.jpain.2019.02.011)] [Medline: [30831273](https://pubmed.ncbi.nlm.nih.gov/30831273/)]
12. Ickmans K, Rheel E, Rezende J, Reis FJ. Spreading the word: pediatric pain education from treatment to prevention. *Arch Physiother* 2022 Nov 15;12(1):25 [FREE Full text] [doi: [10.1186/s40945-022-00151-4](https://doi.org/10.1186/s40945-022-00151-4)] [Medline: [36376953](https://pubmed.ncbi.nlm.nih.gov/36376953/)]
13. Grasaas E, Helseth S, Fegran L, Stinson J, Småstuen M, Lalloo C, et al. App-based intervention among adolescents with persistent pain: a pilot feasibility randomized controlled trial. *Pilot Feasibility Stud* 2022 Jul 27;8(1):158 [FREE Full text] [doi: [10.1186/s40814-022-01113-0](https://doi.org/10.1186/s40814-022-01113-0)] [Medline: [35897086](https://pubmed.ncbi.nlm.nih.gov/35897086/)]
14. Palermo TM, de la Vega R, Murray C, Law E, Zhou C. A digital health psychological intervention (WebMAP Mobile) for children and adolescents with chronic pain: results of a hybrid effectiveness-implementation stepped-wedge cluster randomized trial. *Pain* 2020 Dec 10;161(12):2763-2774 [FREE Full text] [doi: [10.1097/j.pain.0000000000001994](https://doi.org/10.1097/j.pain.0000000000001994)] [Medline: [32658147](https://pubmed.ncbi.nlm.nih.gov/32658147/)]
15. Tudor Car L, Dhinakaran DA, Kyaw BM, Kowatsch T, Joty S, Theng YL, et al. Conversational agents in health care: scoping review and conceptual analysis. *J Med Internet Res* 2020 Aug 07;22(8):e17158 [FREE Full text] [doi: [10.2196/17158](https://doi.org/10.2196/17158)] [Medline: [32763886](https://pubmed.ncbi.nlm.nih.gov/32763886/)]
16. Schachner T, Keller R, V Wangenheim F. Artificial intelligence-based conversational agents for chronic conditions: systematic literature review. *J Med Internet Res* 2020 Sep 14;22(9):e20701 [FREE Full text] [doi: [10.2196/20701](https://doi.org/10.2196/20701)] [Medline: [32924957](https://pubmed.ncbi.nlm.nih.gov/32924957/)]
17. Palanica A, Flaschner P, Thommandram A, Li M, Fossat Y. Physicians' perceptions of chatbots in health care: cross-sectional web-based survey. *J Med Internet Res* 2019 Apr 05;21(4):e12887 [FREE Full text] [doi: [10.2196/12887](https://doi.org/10.2196/12887)] [Medline: [30950796](https://pubmed.ncbi.nlm.nih.gov/30950796/)]

18. Parmar P, Ryu J, Pandya S, Sedoc J, Agarwal S. Health-focused conversational agents in person-centered care: a review of apps. *NPJ Digit Med* 2022 Feb 17;5(1):21 [FREE Full text] [doi: [10.1038/s41746-022-00560-6](https://doi.org/10.1038/s41746-022-00560-6)] [Medline: [35177772](https://pubmed.ncbi.nlm.nih.gov/35177772/)]
19. Sinha C, Cheng AL, Kadaba M. Adherence and engagement with a cognitive behavioral therapy-based conversational agent (Wysa for chronic pain) among adults with chronic pain: survival analysis. *JMIR Form Res* 2022 May 23;6(5):e37302 [FREE Full text] [doi: [10.2196/37302](https://doi.org/10.2196/37302)] [Medline: [35526201](https://pubmed.ncbi.nlm.nih.gov/35526201/)]
20. Hauser-Ulrich S, Künzli H, Meier-Peterhans D, Kowatsch T. A smartphone-based health care chatbot to promote self-management of chronic pain (SELMA): pilot randomized controlled trial. *JMIR Mhealth Uhealth* 2020 Apr 03;8(4):e15806 [FREE Full text] [doi: [10.2196/15806](https://doi.org/10.2196/15806)] [Medline: [32242820](https://pubmed.ncbi.nlm.nih.gov/32242820/)]
21. Or CK, Karsh BT. A systematic review of patient acceptance of consumer health information technology. *J Am Med Inform Assoc* 2009;16(4):550-560 [FREE Full text] [doi: [10.1197/jamia.M2888](https://doi.org/10.1197/jamia.M2888)] [Medline: [19390112](https://pubmed.ncbi.nlm.nih.gov/19390112/)]
22. Ireland D, Andrews N. Pain ROADMAP: a mobile platform to support activity pacing for chronic pain. *Stud Health Technol Inform* 2019 Aug 08;266:89-94 [doi: [10.3233/SHIT190778](https://doi.org/10.3233/SHIT190778)] [Medline: [31397307](https://pubmed.ncbi.nlm.nih.gov/31397307/)]
23. Andrews NE, Ireland D, Deen M, Varnfield M. Clinical utility of a mHealth assisted intervention for activity modulation in chronic pain: the pilot implementation of pain ROADMAP. *Eur J Pain* 2023 Jul 10;27(6):749-765 [doi: [10.1002/ejp.2104](https://doi.org/10.1002/ejp.2104)] [Medline: [36899447](https://pubmed.ncbi.nlm.nih.gov/36899447/)]
24. Cooper A, Ireland D. Designing a chat-bot for non-verbal children on the autism spectrum. *Stud Health Technol Inform* 2018;252:63-68 [Medline: [30040684](https://pubmed.ncbi.nlm.nih.gov/30040684/)]
25. Ireland D, Atay C, Liddle J, Bradford D, Lee H, Rushin O, et al. Hello Harlie: enabling speech monitoring through chat-bot conversations. *Stud Health Technol Inform* 2016;227:55-60 [Medline: [27440289](https://pubmed.ncbi.nlm.nih.gov/27440289/)]
26. Ireland D, Bradford D, Szepe E, Lynch E, Martyn M, Hansen D, et al. Introducing Edna: a trainee chatbot designed to support communication about additional (secondary) genomic findings. *Patient Educ Couns* 2021 Apr;104(4):739-749 [doi: [10.1016/j.pec.2020.11.007](https://doi.org/10.1016/j.pec.2020.11.007)] [Medline: [33234441](https://pubmed.ncbi.nlm.nih.gov/33234441/)]
27. Ireland D, Hassanzadeh H, Tran SN. Sentimental analysis for AIML-based e-health conversational agents. In: Proceedings of the Neural Information Processing 25th International Conference. 2018 Presented at: Neural Information Processing 25th International Conference; December 13-16, 2018; Siem Reap, Cambodia p. 41-51 [doi: [10.1007/978-3-030-04179-3_4](https://doi.org/10.1007/978-3-030-04179-3_4)]
28. Speech-To-Text. Google. URL: <https://cloud.google.com/speech-to-text>
29. Tardif H, Arnold C, Hayes C, Eagar K. Establishment of the Australasian electronic persistent pain outcomes collaboration. *Pain Med* 2017 Jun 01;18(6):1007-1018 [doi: [10.1093/pm/pnw201](https://doi.org/10.1093/pm/pnw201)] [Medline: [27524828](https://pubmed.ncbi.nlm.nih.gov/27524828/)]
30. Lord SM, Tardif HP, Kepreotes EA, Blanchard M, Eagar K. The Paediatric electronic Persistent Pain Outcomes Collaboration (PaedePPOC): establishment of a binational system for benchmarking children's persistent pain services. *Pain* 2019 Jul 1;160(7):1572-1585 [doi: [10.1097/j.pain.0000000000001548](https://doi.org/10.1097/j.pain.0000000000001548)] [Medline: [30839427](https://pubmed.ncbi.nlm.nih.gov/30839427/)]
31. D'arcy-Sharpe AM. Mindfulness for chronic pain: a comprehensive guide. Pathways Health Ltd. 2020. URL: <https://www.pathways.health/blog/mindfulness-for-chronic-pain/>
32. Smiling Mind. 2023. URL: <https://www.smilingmind.com.au/>
33. Wilson Van Voorhis CR, Morgan BL. Understanding power and rules of thumb for determining sample sizes. *Tutor Quant Methods Psychol* 2007 Sep 01;3(2):43-50 [doi: [10.20982/tqmp.03.2.p043](https://doi.org/10.20982/tqmp.03.2.p043)]
34. Abd-Alrazaq A, Safi Z, Alajlani M, Warren J, Househ M, Denecke K. Technical metrics used to evaluate health care chatbots: scoping review. *J Med Internet Res* 2020 Jun 05;22(6):e18301 [FREE Full text] [doi: [10.2196/18301](https://doi.org/10.2196/18301)] [Medline: [32442157](https://pubmed.ncbi.nlm.nih.gov/32442157/)]
35. Nadal C, Sas C, Doherty G. Technology acceptance in mobile health: scoping review of definitions, models, and measurement. *J Med Internet Res* 2020 Jul 06;22(7):e17256 [FREE Full text] [doi: [10.2196/17256](https://doi.org/10.2196/17256)] [Medline: [32628122](https://pubmed.ncbi.nlm.nih.gov/32628122/)]
36. Guest G, MacQueen KM, Namey EE. Applied Thematic Analysis. Thousand Oaks, CA: SAGE Publications; 2012.
37. Andrews NE, Strong J, Meredith PJ, Gordon K, Bagraith KS. "It's very hard to change yourself": an exploration of overactivity in people with chronic pain using interpretative phenomenological analysis. *Pain* 2015 Jul;156(7):1215-1231 [doi: [10.1097/j.pain.000000000000161](https://doi.org/10.1097/j.pain.000000000000161)] [Medline: [25887466](https://pubmed.ncbi.nlm.nih.gov/25887466/)]
38. Johnston V, Brakenridge C, Valiant D, Ling CL, Andrews N, Gane EM, et al. Using framework analysis to understand multiple stakeholders' views of vocational rehabilitation following acquired brain injury. *Brain Impair* 2022 Nov 23;24(2):347-370 [doi: [10.1017/BrImp.2022.27](https://doi.org/10.1017/BrImp.2022.27)]
39. Vaezipour A, Andrews N, Oviedo-Trespacios O, Amershi F, Horswill M, Johnston V, et al. Exploring driving behaviour from the perspectives of individuals with chronic pain and health professionals. *Appl Ergon* 2022 Jul;102:103755 [doi: [10.1016/j.apergo.2022.103755](https://doi.org/10.1016/j.apergo.2022.103755)] [Medline: [35381464](https://pubmed.ncbi.nlm.nih.gov/35381464/)]
40. Chen H, Cohen P, Chen S. How big is a big odds ratio? Interpreting the magnitudes of odds ratios in epidemiological studies. *Commun Stat Simul Comput* 2010 Mar 31;39(4):860-864 [doi: [10.1080/03610911003650383](https://doi.org/10.1080/03610911003650383)]
41. Chan A, Kow R, Cheng JK. Adolescents' perceptions on smartphone applications (Apps) for health management. *J Mobile Technol Med* 2017 Aug;6(2):47-55 [doi: [10.7309/jmtm.6.2.6](https://doi.org/10.7309/jmtm.6.2.6)]
42. Pitardi V, Bartikowski B, Osburg VS, Yoganathan V. Effects of gender congruity in human-robot service interactions: the moderating role of masculinity. *Int J Inf Manage* 2023 Jun;70:102489 [doi: [10.1016/j.ijinfomgt.2022.102489](https://doi.org/10.1016/j.ijinfomgt.2022.102489)]
43. Shumanov M, Johnson L. Making conversations with chatbots more personalized. *Comput Hum Behav* 2021 Apr;117:106627 [doi: [10.1016/j.chb.2020.106627](https://doi.org/10.1016/j.chb.2020.106627)]

44. Chaves AP, Gerosa MA. How should my chatbot interact? A survey on social characteristics in human–chatbot interaction design. *Int J Hum Comput Interact* 2020 Nov 08;37(8):729-758 [doi: [10.1080/10447318.2020.1841438](https://doi.org/10.1080/10447318.2020.1841438)]
45. Kids and family reading report™. 7th edition. Finding their story. Scholastic Inc. 2019. URL: <https://eric.ed.gov/?id=ED594490> [accessed 2023-03-14]
46. Kinney M, Seider J, Beaty AF, Coughlin K, Dyal M, Clewley D. The impact of therapeutic alliance in physical therapy for chronic musculoskeletal pain: a systematic review of the literature. *Physiother Theory Pract* 2020 Aug 28;36(8):886-898 [doi: [10.1080/09593985.2018.1516015](https://doi.org/10.1080/09593985.2018.1516015)] [Medline: [30265840](https://pubmed.ncbi.nlm.nih.gov/30265840/)]
47. Buchman DZ, Ho A, Illes J. You present like a drug addict: patient and clinician perspectives on trust and trustworthiness in chronic pain management. *Pain Med* 2016 Aug 11;17(8):1394-1406 [FREE Full text] [doi: [10.1093/pm/pnv083](https://doi.org/10.1093/pm/pnv083)] [Medline: [26759389](https://pubmed.ncbi.nlm.nih.gov/26759389/)]
48. Irwin CEJ. Using technology to improve the health and well-being of adolescents and young adults. *J Adolesc Health* 2020 Aug;67(2):147-148 [FREE Full text] [doi: [10.1016/j.jadohealth.2020.05.019](https://doi.org/10.1016/j.jadohealth.2020.05.019)] [Medline: [32739018](https://pubmed.ncbi.nlm.nih.gov/32739018/)]
49. Giovanelli A, Ozer EM, Dahl RE. Leveraging technology to improve health in adolescence: a developmental science perspective. *J Adolesc Health* 2020 Aug;67(2S):S7-13 [FREE Full text] [doi: [10.1016/j.jadohealth.2020.02.020](https://doi.org/10.1016/j.jadohealth.2020.02.020)] [Medline: [32718517](https://pubmed.ncbi.nlm.nih.gov/32718517/)]
50. Beaudry J, Consigli A, Clark C, Robinson KJ. Getting ready for adult healthcare: designing a chatbot to coach adolescents with special health needs through the transitions of care. *J Pediatr Nurs* 2019 Nov;49:85-91 [doi: [10.1016/j.pedn.2019.09.004](https://doi.org/10.1016/j.pedn.2019.09.004)] [Medline: [31644960](https://pubmed.ncbi.nlm.nih.gov/31644960/)]
51. Shah J, DePietro B, D'Adamo L, Firebaugh ML, Laing O, Fowler LA, et al. Development and usability testing of a chatbot to promote mental health services use among individuals with eating disorders following screening. *Int J Eat Disord* 2022 Sep 18;55(9):1229-1244 [FREE Full text] [doi: [10.1002/eat.23798](https://doi.org/10.1002/eat.23798)] [Medline: [36056648](https://pubmed.ncbi.nlm.nih.gov/36056648/)]
52. Vigilante VA, Hossain J, Wysocki T, Sharif I. Correlates of type and quantity of child communication during pediatric subspecialty encounters. *Patient Educ Couns* 2015 Nov;98(11):1352-1359 [doi: [10.1016/j.pec.2015.05.001](https://doi.org/10.1016/j.pec.2015.05.001)] [Medline: [26160039](https://pubmed.ncbi.nlm.nih.gov/26160039/)]

Abbreviations

CSIRO: Commonwealth Scientific and Industrial Research Organisation

ePPOC: electronic Persistent Pain Outcomes Collaboration

Edited by A Mavragani; submitted 14.03.23; peer-reviewed by R Marshall, N Pope; comments to author 17.08.23; revised version received 28.08.23; accepted 28.08.23; published 06.10.23

Please cite as:

Andrews NE, Ireland D, Vijayakumar P, Burvill L, Hay E, Westerman D, Rose T, Schlumpf M, Strong J, Claus A
Acceptability of a Pain History Assessment and Education Chatbot (Dolores) Across Age Groups in Populations With Chronic Pain: Development and Pilot Testing
JMIR Form Res 2023;7:e47267
URL: <https://formative.jmir.org/2023/1/e47267>
doi: [10.2196/47267](https://doi.org/10.2196/47267)
PMID: [37801342](https://pubmed.ncbi.nlm.nih.gov/37801342/)

©Nicole Emma Andrews, David Ireland, Pranavie Vijayakumar, Lyza Burvill, Elizabeth Hay, Daria Westerman, Tanya Rose, Mikaela Schlumpf, Jenny Strong, Andrew Claus. Originally published in JMIR Formative Research (<https://formative.jmir.org>), 06.10.2023. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Formative Research, is properly cited. The complete bibliographic information, a link to the original publication on <https://formative.jmir.org>, as well as this copyright and license information must be included.