in Digital Health

Conversational Physical Activity Coaches for Spanish and English Speaking Women: A User Design Study

Caroline A. Figueroa[1]*, Tiffany C. Luo[1], Andrea Jacobo[2], Alan Munoz[1], Minx Manuel[2], David Chan[3], John Canny[3] and Adrian Aguilera[1,4]

[1] School of Social Welfare, University of California, Berkeley, Berkeley, CA, United States, [2] School of Public Health, University of California, Berkeley, Berkeley, CA, United States, [3] Department of Electrical Engineering and Computer Sciences, University of California, Berkeley, Berkeley, CA, United States, [4] Department of Psychiatry and Behavioral Sciences, Zuckerberg San Francisco General Hospital, University of California, San Francisco, San Francisco, CA, United States

Introduction: Digital technologies, including text messaging and mobile phone apps, can be leveraged to increase people's physical activity and manage health. Chatbots, powered by artificial intelligence, can automatically interact with individuals through natural conversation. They may be more engaging than one-way messaging interventions. To our knowledge, physical activity chatbots have not been developed with low-income participants, nor in Spanish—the second most dominant language in the U.S. We recommend best practices for physical activity chatbots in English and Spanish for low-income women.

Methods: We designed a prototype physical activity text-message based conversational agent based on various psychotherapeutic techniques. We recruited participants through SNAP-Ed (Supplemental Nutrition Assistance Program Education) in California (Alameda County) and Tennessee (Shelby County). We conducted qualitative interviews with participants during testing of our prototype chatbot, held a Wizard of Oz study, and facilitated a co-design workshop in Spanish with a subset of our participants.

Results: We included 10 Spanish- and 8 English-speaking women between 27 and 41 years old. The majority was Hispanic/Latina ($n = 14$), 2 were White and 2 were Black/African American. More than half were monolingual Spanish speakers, and the majority was born outside the US ($>50\%$ in Mexico). Most participants were unfamiliar with chatbots and were initially skeptical. After testing our prototype, most users felt positively about health chatbots. They desired a personalized chatbot that addresses their concerns about privacy, and stressed the need for a comprehensive system to also aid with nutrition, health information, stress, and involve family members. Differences between English and monolingual Spanish speakers were found mostly in exercise app use, digital literacy, and the wish for family inclusion.

Conclusion: Low-income Spanish- and English-speaking women are interested in using chatbots to improve their physical activity and other health related aspects. Researchers developing health chatbots for this population should focus on issues

of digital literacy, app familiarity, linguistic and cultural issues, privacy concerns, and personalization. Designing and testing this intervention for and with this group using co-creation techniques and involving community partners will increase the probability that it will ultimately be effective.

INTRODUCTION

Insufficient physical activity is one of the leading risk factors of death worldwide (1). Marginalized groups, such as people of lower socioeconomic status (2), women (3), and ethnic/racial minorities including Latinas (4), are particularly inactive. We need to develop interventions that help marginalized populations increase and maintain healthy physical activity behaviors.

In recent years, digital behavioral health interventions, using smartphones and text-messaging, have seen a surge in development and use (5). These tools can make interventions more accessible: people across a wide range of socioeconomic groups own smartphones in the US, and ownership continues to increase globally (6). The percentage of Latinx individuals that own smartphones is currently 85%, comparable to Blacks (83%) and non-Hispanic whites (85%), and slightly higher than low-income Americans (76%) (7). Although ownership rates are high, disparities in digital literacy and data plan limits persist for ethnic minority and lower-income individuals (8). These factors necessitate user friendly and low data solutions.

The field of digital health has seen an increase in interest in conversational agents, or chatbots, to help individuals pursue healthy lifestyles (9). Chatbots for behavior change can inform and educate, check and monitor symptoms, and improve mental health through Cognitive Behavioral Therapy (10), motivational interviewing, and other therapeutic modalities (11). Chatbots can save costs, guard anonymity, and personalize content (12). Since chatbots can communicate with text-message or voice dialogue, lower digitally skilled users can easily engage (13). Finally, chatbots using natural language processing may be more powerful than one-way messaging because individuals can interact through a natural form of communication (14).

However, most health chatbots may be unsuitable for lower-income and ethnic minority individuals including Latinxs. Most health promotion chatbots offer only English as a communication language (15) despite over 40 million people in the U.S. speaking Spanish at home (16). For example, a systematic review on health chatbots (including general health and mental health) found that out of 45 chatbot studies, only one study used a Spanish chatbot (17). Similarly, our previous review on conversational agents for physical activity identified only a handful of conversational agents for mobile delivery through apps and text messaging, and no Spanish conversational agents (15). Further, digital solutions are not often specifically designed for women (18), especially ethnic minority women, who generally have worse health outcomes than white women (19). Ethnic minority women are underrepresented in the design and testing of digital health tools (18). Because of this and the lack of adequate digital skills training, marginalized groups may underuse these interventions (20), even if health technologies target them.

Thus, developing health chatbots for marginalized populations, including low-income women and Spanish speakers, is challenging but crucial and has not been attempted often enough. We developed a prototype text-message based physical activity conversational agent for low-income English and Spanish speaking individuals, based on principles of Behavioral Activation (21), Motivational Interviewing (22), Acceptance and Commitment Therapy (23), and Solution-Focused Brief Therapy (24). We designed the prototype chatbot to help users clarify their values and motivations for physical activity, set goals and plans, and overcome exercise barriers.
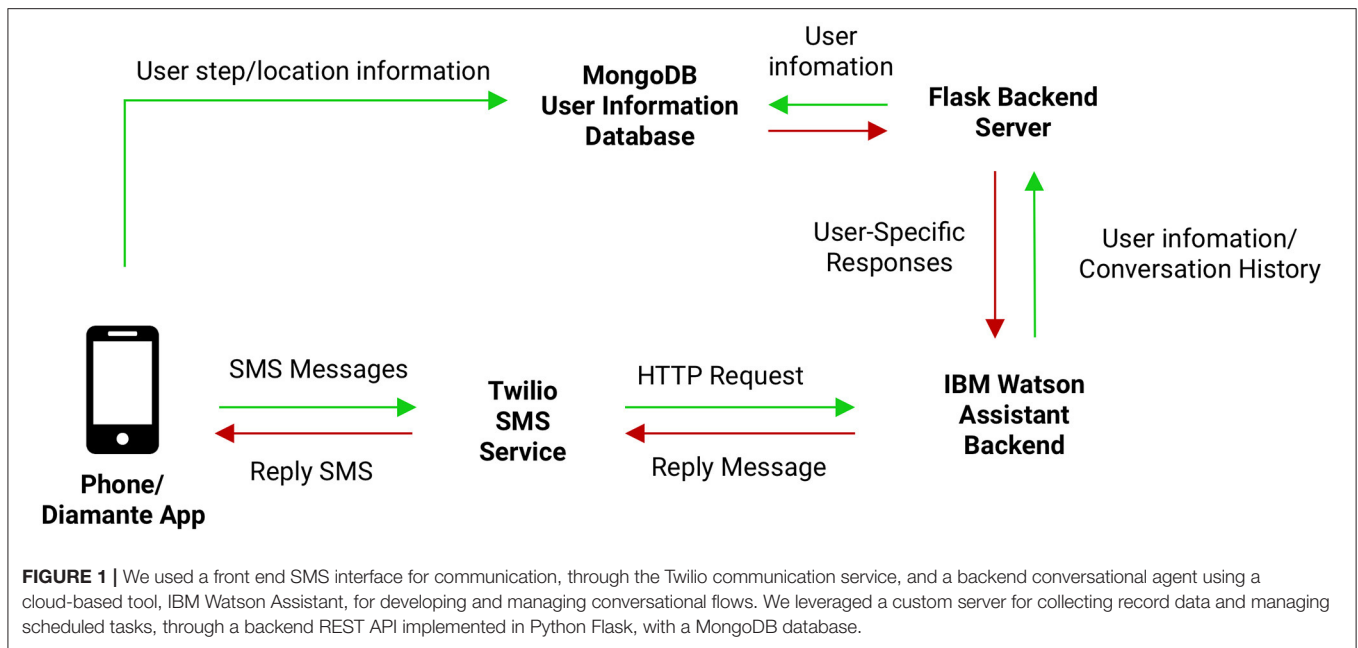
The aim of this study was to understand whether low-income English and Spanish speakers want to interact with conversational agents, what their health priorities are, and how we should design these tools to meet their needs. We report the results of qualitative interviews in low-income Spanish and English speaking women, majority Latinas, during testing of our prototype chatbot, a Wizard of Oz study, and a co-design workshop conducted in Spanish. We originally opened our study to all genders, but received interest only from participants who identified as female. We recommend best practices for researchers interested in developing chatbots for low-income women in English and Spanish.

MATERIALS AND METHODS

Participants

We recruited participants through SNAP-Ed (Supplemental Nutrition Assistance Program Education) in California (Alameda County) and Tennessee (Shelby County). We consider participants low-income as the intended audience for SNAP-Ed is SNAP recipients (who are at or below 200% Federal poverty guidelines) and other low-income audiences who are at or below 185% Federal poverty guidelines.[1] SNAP-Ed extension partners distributed English and Spanish flyers and invited participants in their healthy eating groups to join the study. In addition, SNAP-Ed posted our English and Spanish flyers on their Facebook pages. We included participants 18–65 years old, English or Spanish speaking, who owned a mobile phone and desired to be more physically active. This study was approved by the University of California Berkeley Committee for Protection of Human Subjects (CPHS, ref: 2020-05-13271).

---

[1]https://www.fns.usda.gov/snap/snap-ed

FIGURE 1 | We used a front end SMS interface for communication, through the Twilio communication service, and a backend conversational agent using a cloud-based tool, IBM Watson Assistant, for developing and managing conversational flows. We leveraged a custom server for collecting record data and managing scheduled tasks, through a backend REST API implemented in Python Flask, with a MongoDB database.

Conversational Agent Technical Development

We used a front end SMS interface for communication, through the Twilio communication service[2], and a backend conversational agent using a cloud-based tool, IBM Watson Assistant[3], for developing and managing conversational flows. We leveraged a custom server for collecting record data and managing scheduled tasks (see **Figure 1**), through a backend REST API implemented in Python Flask[4], with a MongoDB[5] database. Our dialog tree consists of over 150 conversational elements, with over 10 million possible conversational pathways.

Conversational Agent Flow Development

Our prototype conversational flows went through two iterations during the study. The first iteration included flows on core values, motivation for health behavior change, exercise goals, and activity planning. Midway through the study we added flows on weekly exercise plans and barriers to physical activity, and we addressed errors after receiving participant feedback. Examples of the dialogue flows are shown in the **Supplementary Material**.

Study Measures

Phase 1: Online Interview

We interviewed participants online between August 2020 and April 2021. After participants filled in an online informed consent form, they were directed to a short online questionnaire that assessed gender, age, socioeconomic status, depression, anxiety, perceived stress, and financial issues due to the COVID-19 pandemic. Within a semi-structured interview conducted via videoconferencing, we assessed opinions and knowledge of chatbots as personal health coaches, technology use, digital

literacy, and privacy considerations of chatbots in general. We also asked participants what information and text-messages they would like to receive from chatbots, and how chatbots could help them remain active during social distancing measures.

Phase 2: Wizard of oz Procedure

After answering these exploratory questions, participants completed a 20 min text-messaging conversation with a simulated chatbot. Participants were told that they were texting an automated chatbot but were actually interacting with a second researcher who texted participants via a Google Voice phone number. This procedure allowed us to understand how a chatbot should ideally respond in conversations with humans (25). Participants were debriefed at the end of the interview about the Wizard of Oz procedure.

Phase 3: Chatbot Prototype

Participants texted with our chatbot prototype for 10–20 min. After the texting conversation, participants resumed the semi-structured interview via videoconference. Participants were asked about their opinions of the chatbot prototype's ease of use, usefulness, humanness, and sustainability, as well as their preference for the simulated chatbot vs. the automated chatbot. They were also invited to give recommendations for expanding the chatbot prototype's content and improving upon its usability. The semi-structured interviews were recorded via Zoom and audio-recordings were transcribed by a professional transcription service (3playMedia). Participants received a $40 gift card upon study completion.

Phase 4: Design Workshop

Midway through the study, two researchers facilitated a co-design session in Spanish via Zoom for Spanish-speaking and bilingual participants who had completed the online interview and tested the chatbot prototype. During the workshop, participants added ideas for chatbot use and design to a Google Jamboard

---

[2]https://www.twilio.com/
[3]https://www.ibm.com/cloud/watson-assistant
[4]https://flask.palletsprojects.com/en/1.1.x/
[5]https://www.mongodb.com/cloud

(**Supplementary Material**). Participants were compensated an additional $40 for their participation.

Analysis

We show descriptives of the clinical and demographic characteristics of the sample using means, medians and percentages where appropriate.

We applied a constructivist grounded theory approach (26) to our qualitative analysis of the semi-structured interviews. TCL developed an initial codebook in English, which AM and AJ used to develop a Spanish codebook. We also employed open coding to generate codes inductively. After all co-authors met to discuss emerging themes, we revised our codebook to incorporate inductively- and deductively derived codes that captured user feedback and attitudes related to the content and physical activity motivators. The transcripts were coded independently by TCL, MM, AJ, and AM using Dedoose, a qualitative software program (27).

The codebook was divided into two distinct parts corresponding to Phases 1 and 3 of the study. The first part of the codebook was designed to analyze participants' comfort level with technology prior to engaging with the simulated chatbot and our chatbot prototype. Data from Phase 1 of the study were tagged with the following codes: "Technology familiarity" and "Privacy and security concerns." The second part of the codebook synthesized data about participants' opinions of our prototype chatbot after engaging in a 10–20 min texting conversation with it. Data from Phase 3 of the study were tagged with the following codes: "Ease of use," "Satisfaction, usefulness, and humanness," "Sustainability," "Physical activity barriers," "Content and usability recommendations." We memoed on the data tagged with these codes to generate the themes we present in this paper.

RESULTS

Participants

We included 10 Spanish- and 8 English-speaking women aged 27–41 years. Most participants ($n = 16$) were recruited by SNAP-Ed Educators. The majority were Hispanic/Latina ($n = 14$), 2 were White and 2 were Black/African-American. More than half were monolingual Spanish speakers, and the majority were born outside the US (>50% in Mexico). For most, paying for basics was hard or somewhat hard ($n = 17$). Over half of the participants had finished high school and/or college ($n = 11$). Most participants ($n = 16$) reported being in good health, and depression and anxiety scores were low overall (>2 indicates risk for depression or anxiety). **Table 1** shows participant characteristics.

Qualitative Findings

We divided the qualitative findings from the semi-structured interviews in Phases 1 and 3 into English and Spanish speaking participant feedback. Here we describe the main results. Additional quotes from participants can be found in the **Supplementary Material**.

TABLE 1 | Participant characteristics.

	Overall ($n = 18$)
Preferred language	**n (%)**
English	8 (44.4%)
Spanish	10 (55.6%)
Age	**Mean (SD)**
	39.5 (6.01)
Race/ethnicity	**n (%)**
Black/African-American	2 (11.1%)
White or Caucasian	2 (11.1%)
Hispanic/Latina	14 (77.8%)
Employment	**n (%)**
Full time (more than or equal to 35 h)	3 (16.7%)
Part time (<35 h)	1 (5.6%)
Homemaker	10 (55.6%)
Unemployed	2 (11.1%)
Retired	1 (5.6%)
Other	1 (5.6%)
Education	**n (%)**
Never went to school	1 (5.6%)
Between 1st and 5th grade	2 (11.1%)
Between 6th and 8th grade	2 (11.1%)
Some high school	2 (11.1%)
High school graduate or "GED" degree	3 (16.7%)
Some college or technical school	1 (5.6%)
College graduate	5 (27.8%)
Graduate degree	2 (11.1%)
Paying for basics (e.g., food, housing, medical care, and heating) is:	**n (%)**
Very hard	4 (22.2%)
Somewhat hard	13 (72.2%)
Not hard at all	1 (5.6%)
Born in the US	**n (%)**
	5 (27.8%)
Country of birth (if not US)	**n (%)**
El Salvador	2 (11.2%)
Guatemala	1 (5.6%)
Mexico	9 (50.7%)
Peru	1 (5.6%)
Self reported-health	**n (%)**
Fair	2 (11.1%)
Good	10 (55.6%)
Very good	3 (16.7%)
Excellent	3 (16.7%)
Impact of COVID-19 (1 = completely disagree, 5= completely agree)	**Mean (SD)**
I am running into financial issues	2.67 (1.37)
I feel more lonely	2.00 (0.970)
I feel more stressed	2.78 (1.48)
I feel more anxious	3.00 (1.41)
Psychological measures	**Mean (SD)**
Depression (PHQ-2, range 0–6)	1.06 (1.06)
Anxiety (GAD-2, range 0–6)	1.67 (1.28)

PHQ-2, Patient Health Questionnaire, 2-item scale; GAD-2, Generalized Anxiety Questionnaire, 2-item scale.

Phase 1 Findings (Prior to Chatbot Conversations)
Technology Familiarity
English-Speaking Participants. Most English-speaking participants had been sending and receiving text (SMS) and multimedia (MMS) messages for years, primarily through their phone's default messaging app but also through other messaging apps (e.g., Facebook Messenger, Google Hangouts, WhatsApp).

> Texting familiarity: "*I want to say over 10 years...15 years I've been texting, forever.*"

Some participants reported also using their mobile phones for health purposes, such as diet and activity tracking (most commonly through Apple's Health app).

> Health apps: "*I did have a fitness app... There's, obviously, the iPhone's health app, in general, for step counting. And my work handed out a health app, and I purchased Zombies, Run!*"

The majority were unfamiliar with the term "chatbot," but said to have encountered a chatbot when the researcher gave examples– e.g., for food ordering and delivery, banking, internet service, and health care.

> Chatbot experience: "*I think I might have [interacted with a chatbot]. Comcast, you know the Help Center? Yeah. I texted them and then I got a response. And then I texted something else, and I got another response. And at the end, they sent me a link to see if that could help.*"

English-speaking participants' initial perceptions of chatbots varied: some were optimistic about the potential uses of chatbots for streamlining interactions between individuals and services (e.g., "*I have no problem with it, especially with knowing the actual human technology that's guiding the technology behind it...I know that artificial intelligence is growing*").

> Others expressed doubt (e.g., "*I know that the world is evolving, and computers are being used for all types of things. But in general, I don't really like it.*").

> Some were skeptical, e.g., "*I think it's going to be frustrating. When I call to make a payment, I do it, and there's no one on the other side. And I ask a question, and [I] get something else.*"

Spanish-Speaking Participants. Besides text-messaging, calling, and social platforms like Facebook's messenger, WhatsApp, Instagram, many Spanish speaking participants also used their phone for basic services.

> Common uses for phones: "*I use it a lot (default texting app), for the bank, for my close family... And the phone...for my basic services: electricity, gas, water, insurance, all that...That's what I use it for more than anything else.*"

Participants used apps like YouTube as a resource for physical health, mental health, and healthy food recipes and followed doctors' advice on Facebook and YouTube.

> Fitness resources: "*I like to watch some videos on YouTube...there are some exercises that I like... They are like boxing with cardio or something... I also like the yoga ones.*"

Most Spanish-speaking participants were unfamiliar with the term "chatbot," until given examples of automated machines (e.g., a phone answering machine).

Even though all had been using text-messages, some participants did not know how to send a new text message to a new number. Additionally, most participants were unfamiliar with video conferencing software and required help from a researcher and their children or partner to navigate Zoom and initiate a conversation with our chatbot.

Privacy and Security Concerns
English-Speaking Participants. Many English-speaking participants were concerned about privacy and were hesitant to share personal information such as their name, images of their face, and location. They were concerned about sharing their location and the chatbot recommending unsafe locations.

> Location recommendations: "*They [the chatbot] say you can go to this park.. but I'm going to be concerned if the park is not being cleaned often, or who gets there. Who goes there? And even though the chatbot is going to tell me this place is safe for you to go, I'm not sure if I'm going to give it a try.*"

Spanish-Speaking Participants. Privacy and security concerns mostly stemmed from limited understanding of chatbot technology. Participants' questions included: "*Do people see what the chatbot is doing? Do people review the messages? Can the chatbot see what I am doing?*"

Additionally, participants were concerned about sharing location data:

> "*I would say that [sharing location] is fine, but on the other hand it would be wrong because I would feel like they know where I am and where I'm going and everything. I would feel like I was being watched.*"

Phase 3 Findings (After Testing Prototype Chatbot)
Ease of Use
English-Speaking Participants. Participants universally appreciated the chatbot's quick response times. As one participant explained, "*the thing I didn't like about the [Wizard of Oz] was that the response time wasn't as fast as the [chatbot].*"

> Others commented that they felt the chatbot "*was going more around the question and not really giving me any solutions.*"

Usability issues were a common complaint. Many participants had to restart the conversation after the chatbot failed to recognize their input.

> Restarting conversations: "*I put in a lot and they sent me back: 'Hi, I am your physical health companion.' So they sent me, again, the same message that they sent me at first.*"

Spanish-Speaking Participants. Participants found the overall experience positive. After the researchers gave coaching and instructions, most participants regardless of literacy level were able to send and receive text messages. After teaching a participant how to create a new message, our team asked the participant if she found the chatbot session to be difficult, to which she replied: *"No, it's easy, I didn't find it difficult myself."*

One participant was unable to reply to the chatbot, as she had difficulties reading, writing and typing. When she needed to send a message, she would use the voice input feature on her phone instead of typing out the message.

Satisfaction, Usefulness, and Humanness

English Speaking Participants. Overall, English-speaking participants were satisfied with their interactions with the chatbot. Participants described the chatbot as *"really great," "pretty conversational," "very responsive,"* and *"very innovative."* One participant stated that the chatbot could be her support person, just like her friend. Some participants preferred texting with the chatbot over texting with a researcher. One participant noted that *"there weren't any grammatical errors in the [chatbot] responses,"* as opposed to the human responses. Another user stated, *"I think this was a little bit better, the second one. Even though I had a little more problems, it was a little more specific or more detailed."*

Participants did not find the chatbot very humanlike, especially when they compared it to their interactions with another researcher via the Wizard of Oz procedure.

> Personalization: *"It seemed like the first one was more conversational–it's more personal, kind of, than the other one. It was just like...answer this question on the second one, to get to the next thing."*

Most participants recognized and accepted the limitations of the chatbot.

> Satisfaction: *"I mean this is cool. Like I said, I would prefer to talk to a live person, but if that's the next best thing, then it's pretty cool."*

Spanish-Speaking Participants. Spanish-speaking participants mentioned the chatbot could give them a quick confidence boost, and help address problems when friends or family members are unavailable. Messages of encouragement and reminders were among the features most liked/requested.

> Encouragement: *"And I get a reminder (from the chatbot) and I need to get back to you (the chatbot) and I can say: 'oh, I am sorry, I am not feeling well today.' And I receive a message saying 'no, it is all right, you can do it!' In other words, a message of encouragement would be very nice."*

Another participant pointed out that humans can get tired of trying to encourage one another, or because of cultural norms, might not be the most supportive.

> Support system: *"Yes, because many times there are people who are very depressed and their relatives don't go to her because they can't go to their house or things like that...that someone (the chatbot) sent them a message I imagine they must get motivated. To say, 'wow, somebody remembered me.'"*

In contrast to the English-speaking participants, most Spanish-speaking participants found their interactions with both the Wizard of Oz chatbot and the automated chatbot to be human-like. One participant stated that she felt like she was texting a person, *"I felt comfortable with the application (chatbot) because it's like I was talking to someone, but it wasn't someone."*

Participants could not differentiate between the researcher and the chatbot and would often ask how the chatbot could be a machine.

Sustainability

English-Speaking Participants. When participants were asked how long they would use the chatbot for, answers ranged from 1 month to indefinitely. Some participants considered engaging with the chatbot daily (and even multiple times throughout the day) for workout ideas, exercise plans, and new exercise knowledge.

> Continued use: *"As long as I'm exercising or if I set a goal for like a month or two months, I think it would be great to have at least reminders or stuff like that, during the whole month, or during these two months. So it would be useful."*

One participant found the chatbot to be more sustainable than an app or activity tracker because *"you're talking to somebody to help you work through things and think through things, and apps aren't really going to do that."*

Some participants felt that the chatbot would lose its novelty over time, especially if its responses became repetitive or unhelpful.

> Repetitiveness: *"If they don't keep up with my goal and they keep asking the same things and not moving on to the place that I am after, probably, two months from now, that will make me stop using it."*

Spanish-Speaking Participants. Most participants stated that they would continue to interact with the chatbot if it remained free. They also shared that they would primarily use it when they need suggestions for healthy eating and new exercises.

> Seeking advice: *"I think when I need help. Like before exercising, ask him[the chatbot] about some exercise...and also about nutrition."*

Participants also mentioned wanting to invite their family to engage with the chatbot, e.g., *"Well, not only me, I think that even the children would like it to motivate them, something like that."*

Physical Activity Barriers

English-Speaking Participants. The majority of our participants expressed a desire to become more physically active, but also shared several barriers that prevent them from achieving their exercise goals. These barriers include work (*"This month I haven't been doing as much [exercise] because I've been a little bit busier*

TABLE 2 | Specific recommendations for content and functions.

Content	1)	Providing pictures and videos of suggested exercises
	2)	Including healthy eating tips
	3)	Sending periodic exercise reminders throughout the day
	4)	Sending daily fitness tips
	5)	Describing health benefits of physical activity in more detail
	6)	Setting milestones for personal fitness goals
	7)	Including weather- and location- specific exercise suggestions (e.g., activities that can be done in the house when it's raining outside)
Functions	1)	Allowing for alternative text input methods (e.g., integrating voice-to-text)
	2)	Allowing users to input their own barriers to physical activity
	3)	Giving users the option to select answers from drop-down lists
	4)	Integrating more opportunities for social connection (e.g., a chat room or integration to existing social platforms)

with work"); lack of companionship and social support (*"It's not motivating to start exercising by myself"*); lack of routine; injury; gym closures; and weather *(e.g., Winter is also coming, so it's like people are going to be inside even more)*.

Participants found that their daily lives were impacted due to caregiving duties. "*I have three girls of different ages and they all need something from me all the time. So I don't have much time for myself. So if I find a time to do something that I like, I just do [my errand] instead of like, stop because the park is on my way.*"

Spanish-Speaking Participants. Most physical activity barriers for the Spanish-speaking group overlapped with English-speaking participants. One participant shared how she was impacted by the stay-at-home-orders.

> Fear of infection: *"I felt very scared because the news came like a bomb…people are dying and they are getting infected, so I already imagined that we were going to get infected…it was like a trauma for me, the truth is that I stopped doing my things from one day to the next."*

For a few participants, physical injuries limited their physical activities.

> Bodily pain: *"I suffer a lot from my lower back and I like to see exercises that help me with that…I really like yoga."* Participants identified personal physical activity recommendations as a benefit of technology like our chatbot prototype.

Content and Usability Recommendations (English and Spanish Participants)
Participants gave specific recommendations to expand upon the program content and chatbot functions. We show the main recommendations in **Table 2**.

Co-design Workshop in Spanish
Background information is shown in the **Supplementary Material**. Here we report the main results.

Participants mentioned that they would like our chatbot to help with activities other than exercise, including finding health information, healthy recipes, COVID-19 vaccine information, and anti-stress tips. If they could create their own ideal chatbot, they would like it to also help them with payment reminders, advice about healthy environments, exercise and cooking tips, and recommendations for visiting new places. Finally, they would design chatbots that help young people study and go to university, and help with parenting, including activities for children.

DISCUSSION

We describe the opinions of low-income English- and Spanish-speaking women on health chatbots, and their experiences with our prototype physical activity chatbot. Overall we find that participants were largely unfamiliar with chatbots, and were initially skeptical of their use. After testing our prototype, most users felt positively about physical activity chatbots. They were concerned about the privacy of their personal information, especially involving location. Users indicated wanting a more comprehensive chatbot system that provides daily exercise goals, tips, health information and healthy food/cooking options. We found differences between English and monolingual Spanish speakers in exercise app use, digital literacy, and the wish for family inclusion. Below we provide specific recommendations for designing health chatbot interventions for this population.

Chatbots, Despite Being Machines, Can Provide Physical Activity Support
Participants were initially skeptical of communicating with a machine. After testing, most commented that the chatbot could give them support and tips when family and friends failed to provide it. Both English- and Spanish-speaking participants found the chatbot to be motivating, similar to a friend that encourages you to exercise. Previous work on automated text-messaging for mental health, also highlighted that Spanish speakers in particular perceive social support from automated text-messages (28). Thus, even if participants know they are talking to a machine, the chatbot could still provide emotional support, particularly when others are too busy or unwilling to provide it.

Pay Attention to Linguistic and/or Cultural Differences Between English and (Mono-Lingual) Spanish Speakers
In contrast to our English-speaking participants, Spanish-speakers used few health apps. Instead, they used social media such as Youtube to find health information. Further, Spanish-speakers generally had lower tech literacy and higher barriers to participating in health tech studies, in line with results from previous studies (8, 29). Several Spanish-speaking participants needed help from the researchers to set up a Zoom meeting and initiate a conversation with the chatbot. For inclusive digital health, this support should be a standard component of the research visit. Finally, though many participants indicated a wish

to add a social component to the chatbot, Spanish speakers in particular wanted their family members (including their children) to be able to engage with the chatbot. Thus, differences were found mostly in exercise app use, digital literacy, and the wish for family inclusion.

Use Simple Text-Based Chatbots, Allow Voice Communication, and Provide Tech Support

Our findings combined with earlier work (30) demonstrates an advantage of a text-based chatbot rather than an app-only interface. Most participants were familiar with texting, but not all were frequent app users. Other work in Latinx parents also suggested that using tools that participants are already familiar with for chatbots, such as Whatsapp will be most effective. For participants with lower reading, writing and tech literacy, voice interactions could further increase the usability of chatbots. For example, a user co-creation study showed that older adults, who generally have lower digital literacy, prefer voice based physical activity chatbots to text based (31). The authors suggested that voice is a powerful modality for encouraging motivation among those who struggle with new technologies.

Participants also indicated they preferred short and easy-to-understand messages. In line with other work (32), we recommend that the reading level should be no more than 8th grade (13–14 years) literacy. Finally, technical issues (e.g., chatbot struggled to understand answers outside of a template) impeded smooth communication. Flexibility of bot response, dialogue length, dialogue structure, and chatbot personality are general technical challenges of chatbot design, which need to be addressed and improved in future work to avoid user frustration and drop-out (33).

Increase Transparency About Data Collection

Most users were concerned about sharing personal information, particularly their location. Users felt more comfortable with data sharing if they understood the reason for that data collection (e.g., giving more personalized tips). Informing users of the reason behind requesting information may make them feel more comfortable and willing to use the technology.

Consider Designing Comprehensive Chatbots to Help With Health and Non-health Activities

Most participants indicated they wanted help with healthy eating and finding reliable health information (including COVID-19 related information). Further, many Spanish speakers suggested chatbots should include the rest of their family into healthy living, and also provide parenting tips. Several participants envisioned chatbots also helping with practical activities such as banking, cooking and parenting. Researchers could consider comprehensive wellness chatbots that can connect individuals to other non-health related services. One drawback of this approach is the more functions a chatbot includes, the more difficult it becomes to design the chatbot, and to measure its effects (34).

Factorial designs such as the multiphase optimization strategy might be helpful in this case to measure the effects of various chatbot components (35).

Consider Adaptive Chatbots to Keep Novelty

Users expressed concerns of boredom with chatbots over time if functions and recommendations remain unchanged. In our previous systematic review, we found that when program content was repetitive, users were more likely to disengage from chatbots (15). Most conversational agents included in our review lacked personalization and only communicated through multiple-choice responses. Studies also neglected to discuss safety and privacy issues, and few conversational agents acted on users' mentions of injuries, pain, or mental health symptoms. To increase engagement, chatbots should be adaptive (36–38), for example change the goals and tips based on participant behavior, and add new recommendations, weather and location specific tips and videos. Some participants also mentioned a wish to include their own exercise barriers, and be connected to others through a chatbot. Options would be to allow for more user input, and add peer support, or a human health coach when the chabot fails to provide adequate guidance.

Partnering With Community Organizations and Co-creation

Partnering with local community partners proved to be essential for recruiting typically underrepresented participants into our study. When SNAP-Ed health trainers recommended our study, participants were much more inclined to participate. Through flyers on the Facebook pages of SNAP-Ed, very few participants enrolled (only 2, both white and non-Spanish speaking). Community partners should not only be involved in the recruitment process, but also in the dissemination of health technology innovations. Our co-creation session brought unexpected participant preferences and wishes, which were useful in developing subsequent versions of our chatbot. Co-creation also helps to prevent "bad design," e.g., designing an intervention our target group won't use (39, 40). Thus, engaging both community partners and participants in the design, testing and dissemination of interventions can increase the likelihood that participants will use and benefit from health chatbot interventions.

Limitations

We included low-income women who had already demonstrated a vested interest in their health by joining the SNAP-Ed healthy nutrition program. Women were paid for their participation in the study, which may have impacted their responses and the outcomes of this study. These findings may therefore not generalize to a less motivated group who will not receive compensation. Further, the number of participants was low. This study must be considered a pilot. We divided our findings based on English- and Spanish-speaking women. However, among the English-speaking

women, half were bilingual Latinas (also spoke Spanish). Cultural differences may therefore be small, though the English-speaking participants may be more acculturated. We observed differences between English- and mono-lingual Spanish-speaking participants mostly related to digital literacy and app use.

Future Steps

This research confirms that participants are interested in talking to computer-based agents about social and physical health issues. Future work should test more finalized chatbot applications in this population in larger studies, with a subsequent user-testing phase of a more finalized version of the application and ultimately in a randomized clinical trial. Most systems are currently too inflexible to personalize care because they use rule-based responses. Increasing personalization, creating chatbots that can respond flexibly, and advancing health equity by reaching marginalized populations should be major goals of future chatbot research. By tailoring this intervention to low-income Spanish- and English-speaking women-who can greatly benefit from mobile health applications but for whom they are often not designed-researchers can contribute to improving health and (digital) health equity.

Conclusion

Low-income Spanish- and English-speaking women are interested in using chatbots to improve their physical activity and general health and feel supported by these tools. Issues that researchers should take into account when designing chatbots for this population include digital literacy, app familiarity, linguistic and cultural issues, privacy concerns and novelty and flexibility. Using co-creation techniques and involving community partners will increase the likelihood that health chatbots will be effective. Future work should focus on personalization, co-creation, and ensuring health equity through digital innovations.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by University of California Berkeley Committee for Protection of Human Subjects (CPHS). The patients/participants provided their written informed consent to participate in this study.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fdgth.2021.747153/full#supplementary-material

REFERENCES

1. Warburton DER, Nicol CW, Bredin SSD. Health benefits of physical activity: the evidence. *CMAJ.* (2006) 174:801–9. doi: 10.1503/cmaj.051351
2. Gidlow C, Johnston L, Crone D, Ellis N, James D. A systemic review of the relationship between socio-economic position and physical activity. *Health Educ J.* (2006) 65:338–67. doi: 10.1177/0017896906069378
3. Troiano RP, Berrigan D, Dodd KW, Mâsse LC, Tilert T, McDowell M. Physical activity in the United States measured by accelerometer. *Med Sci Sports Exerc.* (2008) 40:181–8. doi: 10.1249/mss.0b013e31815a51b3
4. Keller C, Fleury J. Factors related to physical activity in hispanic women. *J Cardiovasc Nurs.* (2006) 21:142–5. doi: 10.1097/00005082-200603000-00012
5. Krauss J, Zheng P, Stewart C, Berman M. Digital health technology for behavior change. In: Rippe JM, editor. *Lifestyle Medicine.* 3rd ed. Boca Raton, FL: CRC Press (2019). doi: 10.1201/9781315201108-25
6. Poushter J. *Smartphone Ownership and Internet Usage Continues to Climb in Emerging Economies.* Pew Research Center (2016). Available online

at: https://www.pewresearch.org/global/2016/02/22/smartphone-ownership-and-internet-usage-continues-to-climb-in-emerging-economies/
7. Pew Research Center. *Demographics of Mobile Device Ownership and Adoption in the United States.* Pew Research Center: Internet, Science & Tech (2021). Available online at: https://www.pewresearch.org/internet/fact-sheet/mobile/
8. Nouri S, Khoong EC, Lyles CR, Karliner L. Addressing equity in telemedicine for chronic disease management during the Covid-19 pandemic. *NEJM Catal Innov Care Deliv.* (2020). doi: 10.1056/CAT.20.0123
9. Kocaballi AB, Berkovsky S, Quiroz JC, Laranjo L, Tong HL, Rezazadegan D, et al. The personalization of conversational agents in health care: systematic review. *J Med Intern Res.* (2019) 21:e15360. doi: 10.2196/15360
10. Fitzpatrick KK, Darcy A, Vierhile M. 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.* (2017) 4:e19. doi: 10.2196/mental.7785
11. Zhang J, Oh YJ, Lange P, Yu Z, Fukuoka Y. Artificial intelligence chatbot behavior change model for designing artificial intelligence chatbots to

promote physical activity and a healthy diet: viewpoint. *J Med Intern Res.* (2020) 22:e22845. doi: 10.2196/22845

12. Pereira J, Díaz O. Using health chatbots for behavior change: a mapping study. *J Med Syst.* (2019) 43:135. doi: 10.1007/s10916-019-1237-1

13. Følstad A, Brandtzaeg P. Chatbots and the new world of HCI. *Interactions.* (2017) 24:38–42. doi: 10.1145/3085558

14. Schuetzler RM, Grimes GM, Giboney JS. The impact of chatbot conversational skill on engagement and perceived humanness. *J Manage Inform Syst.* (2020) 37:875–900. doi: 10.1080/07421222.2020.1790204

15. Luo TC, Aguilera A, Lyles CR, Figueroa CA. Promoting physical activity through conversational agents: mixed methods systematic review. *J Med Intern Res.* (2021) 23:e25486. doi: 10.2196/25486

16. United States Census Bureau. *Language Spoken at Home (Measuring America's People, Places and Economy, p. Table S1601) [American Community Survey].* U.S. Department of Commerce (2019). Available online at: https://data.census.gov/cedsci/table?q=spanish&tid=ACSST1Y2019.S1601

17. Safi Z, Abd-Alrazaq A, Khalifa M, Househ M. Technical aspects of developing chatbots for medical applications: scoping review. *J Med Intern Res.* (2020) 22:e19127. doi: 10.2196/19127

18. Figueroa CA, Luo TC, Aguilera A, Lyles CR. The need for feminist intersectionality in digital health. *Lancet Dig Health.* (2021) 3:e526–33. doi: 10.1016/S2589-7500(21)00118-7

19. Carratala S, Maxwell C. *Fact Sheet: Health Disparities by Race and Ethnicity.* Center for American Progress (2020). Available online at: https://cdn.americanprogress.org/content/uploads/2020/05/06130714/HealthRace-factsheet.pdf

20. Schueller SM, Hunter JF, Figueroa C, Aguilera A. Use of digital mental health for marginalized and underserved populations. *Curr Treat Opt Psychiatry.* (2019) 6:243–55. doi: 10.1007/s40501-019-00181-z

21. Kanter JW, Manos RC, Bowe WM, Baruch DE, Busch AM, Rusch LC. What is behavioral activation? A review of the empirical literature. *Clin Psychol Rev.* (2010) 30:608–20. doi: 10.1016/j.cpr.2010.04.001

22. Rollnick S, Miller W. What is motivational interviewing? *Behav Cogn Psychother.* (1995) 23:325–34. doi: 10.1017/S135246580001643X

23. Hayes SC, Strosahl KD, Bunting K, Twohig M, Wilson KG. What is acceptance and commitment therapy? In: Hayes SC, Strosahl KD, editors. *A Practical Guide to Acceptance and Commitment Therapy.* Berlin; Heidelberg: Springer (2004). p. 3–29. doi: 10.1007/978-0-387-23369-7_1

24. Shazer S, de, Dolan Y, Korman H, Trepper T, McCollum E, et al. *More Than Miracles: The State of the Art of Solution-Focused Brief Therapy.* 2nd edn. New York, NY: Routledge (2021). doi: 10.4324/9781003125600

25. Kelley JF. An empirical methodology for writing user-friendly natural language computer applications. In: *Proceeding of ACM SIG-CHI'83 Human Factors in Computing Systems* (New York, NY). (1983). p. 193–6. doi: 10.1145/800045.801609

26. Mills J, Bonner A, Francis K. Adopting a constructivist approach to grounded theory: implications for research design. *Int J Nurs Pract.* (2006) 12:8–13. doi: 10.1111/j.1440-172X.2006.00543.x

27. Salmona M, Lieber E, Kaczynski D. *Qualitative and Mixed Methods Data Analysis Using Dedoose: A Practical Approach for Research Across the Social Sciences.* 1st edn. Thousand Oaks, CA: Sage Publications (2019).

28. Aguilera A, Berridge C. Qualitative feedback from a text messaging intervention for depression: BENEFITS, drawbacks, cultural differences. *JMIR MHealth UHealth.* (2014) 2:e46. doi: 10.2196/mhealth.3660

29. Hernandez-Ramos R, Aguilera A, Garcia F, Miramontes-Gomez J, Pathak LE, Figueroa CA, et al. Conducting internet-based visits for onboarding populations with limited digital literacy to an mhealth intervention: development of a patient-centered approach. *JMIR Format Res.* (2021) 5:e25299. doi: 10.2196/25299

30. Jain M, Kumar P, Kota R, Patel SN. Evaluating and informing the design of chatbots. In: *Proceedings of the 2018 Designing Interactive Systems Conference* (New York, NY). (2018). p. 895–906. doi: 10.1145/3196709.3196735

31. Wiratunga N, Cooper K, Wijekoon A, Palihawadana C, Mendham V, Reiter E, et al. FitChat: conversational artificial intelligence interventions for encouraging physical activity in older adults. *arXiv preprint arXiv:2004.14067.* (2020).

32. Pathak LE, Aguilera A, Williams JJ, Lyles CR, Hernandez-Ramos R, Miramontes J, et al. Developing messaging content for a physical activity smartphone app tailored to low-income patients: user-centered design and crowdsourcing approach. *JMIR MHealth UHealth.* (2021) 9:e21177. doi: 10.2196/21177

33. Fadhil A, Schiavo G. Designing for health chatbots. *arXiv preprint arXiv:1902.09022.* (2019).

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 Intern Res.* (2020) 22:e18301. doi: 10.2196/18301

35. Collins LM, Murphy SA, Strecher V. The multiphase optimization strategy (MOST) and the sequential multiple assignment randomized trial (SMART): new methods for more potent ehealth interventions. *Am J Prevent Med.* (2007) 32:S112–8. doi: 10.1016/j.amepre.2007.01.022

36. Figueroa CA, Aguilera A, Chakraborty B, Modiri A, Aggarwal J, Deliu N, et al. Adaptive learning algorithms to optimize mobile applications for behavioral health: guidelines for design decisions. *J Am Med Inform Assoc.* (2021) 28:1225–34. doi: 10.1093/jamia/ocab001

37. Aguilera A, Figueroa CA, Hernandez-Ramos R, Sarkar U, Cemballi A, Gomez-Pathak L, et al. mHealth app using machine learning to increase physical activity in diabetes and depression: clinical trial protocol for the DIAMANTE study. *BMJ Open.* (2020) 10:e034723. doi: 10.1136/bmjopen-2019-034723

38. Zhou M, Fukuoka Y, Mintz Y, Goldberg K, Kaminsky P, Flowers E, et al. Evaluating machine learning–based automated personalized daily step goals delivered through a mobile phone app: randomized controlled trial. *JMIR MHealth UHealth.* (2018) 6:e9117. doi: 10.2196/mhealth.9117

39. Leask CF, Sandlund M, Skelton DA, Altenburg TM, Cardon G, Chinapaw MJM, et al. Framework, principles and recommendations for utilising participatory methodologies in the co-creation and evaluation of public health interventions. *Res Involve Engage.* (2019) 5:2. doi: 10.1186/s40900-018-0136-9

40. Nijagal M, Patel D, Lyles C, Liao J, Chehab L, Williams S, et al. Using human centered design to identify opportunities for reducing inequities in perinatal care. *Res Square.* (2021) doi: 10.21203/rs.3.rs-138912/v1