Pilot Study

DIGITAL HEALTH

DIGITAL HEALTH
Volume 10: 1–10

DOI: 10.1177/20552076241283243
journals.sagepub.com/home/dhj
Sage

Navigating healthier beverage consumption in adolescents using the "R-Ma Bot" chatbot: A usability and evaluation study

Jisu Lee[1,2], Hyeonkyeong Lee[1,2] and Hyeyeon Lee[1]

Abstract

Objective: This pilot study aimed to evaluate the usability and effectiveness of a behavior change technique (BCT)-based chatbot developed to promote healthier beverage consumption among adolescents.

Methods: The Read and Manage your health roBot ("R-Ma Bot"), designed with 13 BCTs, was tested with 42 adolescents (13 men, 29 women, mean age 15.0 ± 0.7) for 2 weeks. Usability was assessed after the 2-week intervention using a chatbot usability questionnaire, recruitment, retention, participation, and engagement. Scores above 70 out of 100 were considered high usability. Qualitative data from open-ended questions were collected for evaluation. Effectiveness was measured by changes in knowledge, use and impact of nutrition labels, and weekly consumption of sugar, sodium, and caffeine from carbonated and/or energy drinks before and after the 2-week intervention.

Results: The score of R-Ma Bot's usability averaged 74.7, with participants addressing it useful, friendly, and easy to use, though they suggested improving unnatural conversation flow. All participants engaged with the chatbot for at least 13 out of 14 days, with over half using it daily for the entire period. After intervention, awareness of nutrition labels increased from 64.3% to 92.9%, and nonreaders decreased from 42.9% to 16.7%. Weekly sugar intake from beverages significantly decreased by 60%, from 13.1 ± 20.1 mg to 7.9 ± 12.8 mg.

Conclusions: R-Ma Bot's high usability contributed to high retention and behavioral changes, significantly reduced sugar consumption from beverages and improved awareness of nutrition labels. We suggest integrating strategies that enhance knowledge, motivation, and opportunities through BCTs with youth-friendly design elements in the development of interventions for adolescents.

Keywords

Adolescents, public health, behavior change, artificial intelligence, diet

Submission date: 10 May 2024; Acceptance date: 28 August 2024

Introduction

Approximately one-third of adolescents in Korea consume carbonated drinks three or more times per week, a figure that has been increasing rapidly over the past decade.[1] Carbonated drinks contain high levels of sugar and sodium, and their excessive consumption among adolescents can lead to weight gain, obesity, and the development of chronic diseases.[2] Many carbonated drinks contain caffeine, especially energy drinks, which are popular among

[1]Mo-Im Kim Nursing Research Institute, College of Nursing, Yonsei University, Seoul, South Korea
[2]Institute for Innovation in Digital Healthcare, Yonsei University, Seoul, South Korea

Corresponding author:
Hyeonkyeong Lee, Mo-Im Kim Nursing Research Institute, College of Nursing, Yonsei University, 50 Yonsei-ro, Seodaemun-gu, Seoul 03722, Republic of Korea.
Email: hlee39@yuhs.ac

adolescents because they are perceived as decreasing study-induced fatigue, increasing concentration, and promoting socialization with peers.[3] However, caffeine in beverages has been found to induce not only physical symptoms such as headaches and sleeping problems but also psychological symptoms such as stress, anxiety, and aggressive behavior.[4] Considering the increased consumption of carbonated and/or energy drinks and their negative effects, it is clear that efforts should be made to promote healthier beverage consumption habits in this age group.

The consumption of carbonated and/or energy drinks is generally higher among males, individuals with lower education and socioeconomic status, and those with limited knowledge or awareness of the adverse effects of these beverages.[5,6] In addition, low of knowledge and awareness of nutrition labels and a lack of interest or motivation to change their behavior contribute to the consumption of unhealthy beverages.[7] Furthermore, there is no global mandate to list caffeine content on nutrition labels, which could potentially limit adolescents' ability to make informed decisions regarding beverage consumption. Adolescents from multiethnic backgrounds in South Korea may be particularly susceptible to the negative effects related to the consumption of these beverages due to factors such as language, socioeconomic status, and acculturation.[8] Therefore, an effective health promotion approach that considers their vulnerabilities is necessary.

Adolescents accustomed to maintaining relationships and communicating in the digital environment may benefit from interventions utilizing digital technology.[9,10] A chatbot is a digital tool that has recently been recognized as a viable approach for delivering interventions that require ongoing management and monitoring, such as those aimed at dietary changes.[11] However, interventions lacking a theoretical basis for behavior change have demonstrated limited long-term effectiveness.[12] The COM-B model suggests that behavior change occurs when capability, including knowledge and skills, is coupled with opportunity and motivation.[13] The strategies that facilitate this process are the behavior change techniques (BCTs) applied in this study.[14] These strategies have been used as a theoretical framework in health promotion interventions aimed at improving dietary behaviors[15,16] and are also suitable for adolescents.[17] Specifically, effective changes have been observed when integrating "feedback" and "goal-setting" strategies with the capability-enhancing strategies such as "information about health outcomes."[15,18]

Usability is a measure of how effectively and satisfactorily users can achieve their goals.[19] By conducting usability assessments, we can enhance user experience and satisfaction, which in turn positively influences outcomes like feasibility and sustainability. Therefore, evaluating interventions is crucial for developing, improving, and maintaining new interventions. The primary aim of this study was to evaluate the usability and effectiveness of a chatbot-based behavior change intervention for navigating healthier beverage consumption among adolescents. The study's secondary aim was to assess whether the changes in adolescents' beverage consumption behaviors over the 2-week intervention differed by participants' demographic characteristics or participation.

Methods

Study design

This study was a pilot study to determine usability and effectiveness, employing one-group pretest–posttest design. The intervention duration was determined to be 2 weeks (14 days) based on the previous study that demonstrated the effectiveness of chatbot interventions in changing dietary habits.[16] This study used a convergent parallel mixed methods approach to integrate quantitative and qualitative data[20] to provide a comprehensive understanding of how effectively and satisfactorily users can interact with the intervention.

Participants

A sample of 42 adolescents residing in Seoul or Incheon, South Korea, was recruited for this study. The sample size was determined by calculating a two-tailed test with an effect size of 0.73,[21] a significance level of 0.05, a power of 0.80, and an estimated dropout rate of 20%[22] in G-Power 3.1.9. We recruited 21 native Korean adolescents and 21 racial and ethnic adolescents to compare the outcomes of native Korean and racial and ethnic adolescents.

Two middle schools in Seoul and Incheon, South Korea, that agreed to participate in the study and recruit participants were conveniently selected. Recruitment flyers were posted with permission. The researchers explained the study to the students and their parents, who voluntarily expressed interest in participating and obtained informed consent. Recruitment was terminated when each group reached the target number of participants. The inclusion criteria were adolescents aged between 14 and 16 years who could read, write, and speak Korean. Racial and ethnic adolescents were defined as children of multiethnic families in which at least one parent had a foreign nationality. Native adolescents were of Korean nationality, and both parents were born in Korea and had Korean nationality. The exclusion criteria were adolescents with medical conditions requiring dietary therapy or medication.

Intervention

The development process of the Read and Manage your Health (R-Ma) Bot was adopted and modified based on the chatbot life cycle.[23] This modified process was comprised of six stages—defining user and chatbot objectives, analyzing user characteristics and BCTs, composing scripts (i.e., algorithms) based on acquired knowledge, developing possible dialog scenarios, implementing the chatbot, and conducting implementation testing and modifications. All areas of the chatbot, including its algorithms, scenarios, and feedback, were designed by incorporating the 13 BCTs proven in previous studies.[17] The selected BCTs included strategies such as feedback and rewards for behaviors and outcomes, and support systems. Particularly, the information provision stages were subdivided into three levels, and customized capability-enhancing strategies were applied according to the change of individual intake or behavior changes.

The main functions of the Read and Manage your health roBot (R-Ma Bot) included analyzing beverage consumption data and providing feedback based on users' beverage intake. The R-Ma Bot provided a reminder alarm at 20:00 every day to ensure the data is entered daily. It also had a search function for beverage ingredients that allowed participants to explore information about beverages independently. Natural language processing and machine learning were applied to generate flexible responses based on the conversation patterns of the adolescents analyzed in the user analysis step. We constructed dialogue scenarios by considering the users' conversational flow, comprehension level, and engagement potential. The R-Ma Bot was delivered using "K" Company, the most easily accessible social networking service in Korea (Figure 1).

Measures

Usability. The usability of the R-MA Bot was assessed through a semi-structured questionnaire, and also through feasibility metrics such as retention, participation rate, and engagement.

The semistructured questionnaire collected both quantitative and qualitative data. For the quantitative data, we used the Chatbot Usability Questionnaire (CUQ),[24] a 16-item questionnaire consisting of eight positive and eight negative questions. Each item was measured on a 5-point Likert scale, with 1 indicating "*strongly disagree*" and 5 indicating "*strongly agree*." Scores were converted to a 100-point scale using a provided formula, with higher scores indicating greater usability. Although a cut-off point was not explicitly provided, a score of 70 or higher indicated good usability. With the permission from the original author, two experts independently translated the forward and backward translations and reviewed them in a meeting with another expert to reach a consensus before using the Korean version of CUQ. All three experts were fluent in Korean and English and had extensive experience in adolescent health behavior change research. A Korean expert reviewed the translated instrument to ensure that it read naturally. The original tool has been validated for construct validity and reliability,[25] and Cronbach's alpha for the translated instrument was .873 in this study. Qualitative usability data were collected using an open-ended question: "Please describe what you liked about using the chatbot R-Ma Bot and what could be improved."

Retention was defined and calculated as the percentage of participants who responded by the last day of the intervention. The participation rate was calculated as the percentage of participants who entered the data. The type of engagement was classified as active or passive based on whether the data was entered before or after the daily reminder from the R-Ma Bot at 20:00. These measurements were validated by cross verifying the participants' attendance records with their data entry records.

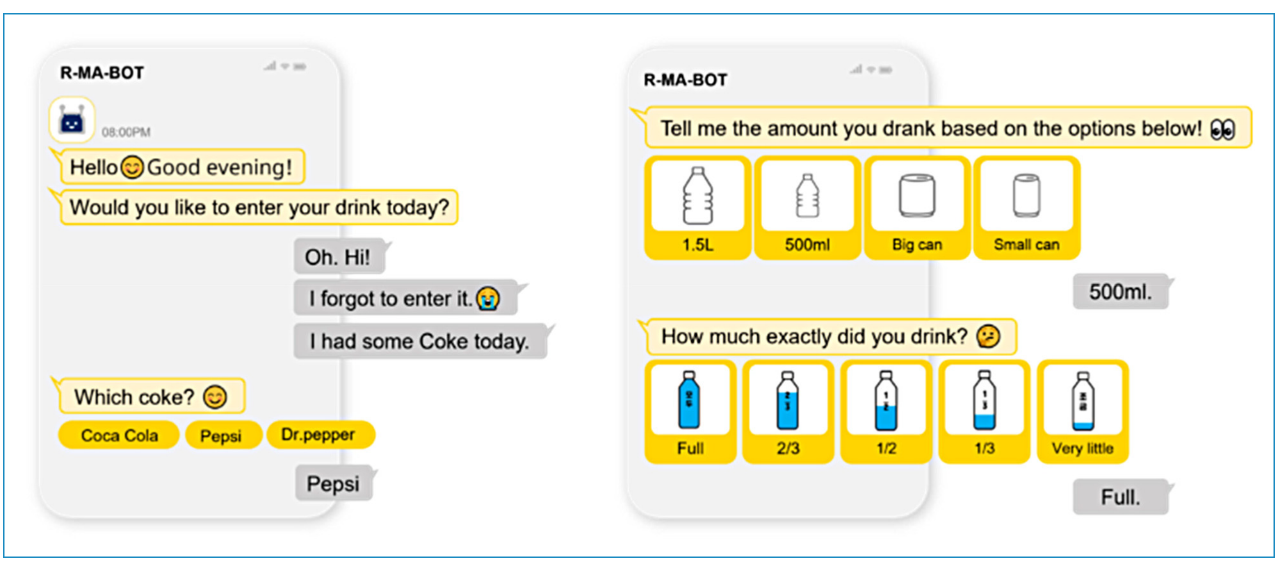

Figure 1. The example dialog in R-Ma Bot.
R-Ma Bot: Read and Manage your health roBot.

Effectiveness. The effectiveness of the intervention was assessed by changes in knowledge and behavior through three questions about changes in knowledge, use and impact of nutrition labels on beverages: "Do you know what a nutrition label is?," "Do you read nutrition labels when buying or choosing a beverage?," and "Do nutrition labels influence your beverage choices?." All the items were adapted from the Korea National Health and Nutrition Examination Survey[26] and were assessed using a yes/no dichotomous scale. The average weekly amounts of sugar, sodium, and caffeine consumed from carbonated and/or energy beverages were assessed and compared between Weeks 1 and 2 to evaluate whether there was a change in consumption. Intake was calculated based on the beverage information provided by the Korean Ministry of Food and Drug Safety, the U.S. Department of Agriculture, and a related website.[27]

Participants' characteristics. These included age, gender, parents' country of birth, academic performance, social status, height, and weight, collected using a self-reported questionnaire. Academic performance over the past 1 year was categorized as high, upper-middle, middle, lower-middle, and low. Scores of middle or above were classified as high, while the rest were classified as low. Social status was assessed using youth version of MacArthur scale of Subjective Social Status, with scores ranging from 1 (*lowest*) to 10 (*highest*).[28] Scores above 5 were classified as middle or higher.

Data collection

Data were collected in January 2023 using a combination of paper and online methods. The pretest was conducted after informed consent from both children and parents and included a written questionnaire on sociodemographic characteristics and knowledge, use, and impact of nutrition labeling. During the 14-day intervention period, participants entered the type and amount of beverages they consumed daily into R-Ma Bot. After the intervention ended, a post-test was administered via an online questionnaire, with usability items added to the pre-test. Pre- and posttest data collection took place for 1 week before and after the 2-week intervention.

Data analysis

Quantitative data was analyzed using descriptive statistics, including frequency, proportion, mean, and standard deviation. In the usability assessment, CUQ scores were presented as mean and standard deviation after undergoing reverse scoring procedures. Qualitative data were analyzed and interpreted using the procedures suggested by Cresswell and Clark (2017).[20] The transcribed data were thoroughly reviewed, and participants' perceptions of what they liked and what they believed needed improvement were categorized. Similar responses were grouped together, and themes were derived through a systematic coding process. During this process, the grouped responses were aligned with the items of the CUQ to ensure that the themes were firmly grounded in the data.

To evaluate effectiveness, changes in knowledge, use, and impact of nutrition labels were analyzed using McNemar's test for paired sample, employing a binominal test. Changes in pre- and postconsumption amounts were analyzed using paired *t*-tests, with additional verification for differences between native and racial and ethnic groups.

Ethics

Ethics approval was obtained from the appropriate Institutional Review Board. As the participants in this study were adolescents, written informed consent was obtained from both study participants and their legally authorized representatives prior to the initiation of the study. The researcher obtained the contact details of the participants' parents or legally authorized representatives from the completed consent forms. Then, we directly contacted them to explain the study and requested written consent.

Results

Sociodemographic characteristics

The mean age was 15.0 years ($SD = 0.7$), and 69.0% ($n = 29$) were women. Among the 21 racial and ethnic youth, six had only mothers born abroad, while 15 had both parents born abroad. The parent countries included Bangladesh, China, Kazakhstan, Mali, Myanmar, Nigeria, Pakistan, Philippines, Russia, Ukraine, Uzbekistan, and Vietnam. The father country was Kazakhstan, and the mother country was China, which accounted for the highest percentage of the racial and ethnic group (Table 1).

Usability: Chatbot usability questionnaire

The mean CUQ score for all participants was 74.7 ($SD = 11.9$; Table 2). Scores ranged from 54.5 to 100.0. There was no difference between the scores of native and racial and ethnic adolescents.

A total of 27 participants provided qualitative responses regarding R-Ma Bot. There were 22 positive responses, including four participants who responded that all parts were good. The themes that emerged from positive responses were "Usefulness," "Friendliness," and "Easy to use," which were mentioned nine, eight, and five times, respectively. However, the five negative responses that could be improved, were the themes of "Unnatural conversation" and "Need more information" (Box 1).

Table 1. Sociodemographic characteristics of the study participants ($N = 42$).

Characteristics		*n* (%)	*M* (*SD*)
Age (years)	14	10 (23.8)	15.0 (0.7)
	15	20 (47.6)	
	16	12 (28.6)	
Gender	Men	13 (31.0)	
	Women	29 (69.0)	
Father nationality	Bangladesh	1 (2.4)	
	China	2 (4.8)	
	Kazakhstan	3 (7.1)	
	Mali	1 (2.4)	
	Myanmar	1 (2.4)	
	Nigeria	1 (2.4)	
	Pakistan	1 (2.4)	
	Philippines	1 (2.4)	
	Russia	1 (2.4)	
	South Korea	27 (64.3)	
	Ukraine	1 (2.4)	
	Uzbekistan	2 (4.8)	
Mother nationality	Bangladesh	1 (2.4)	
	China	5 (11.9)	
	Kazakhstan	2 (4.8)	
	Mali	1 (2.4)	
	Myanmar	1 (2.4)	
	Nigeria	1 (2.4)	
	Pakistan	1 (2.4)	
	Philippines	2 (4.8)	
	Russia	2 (4.8)	
	South Korea	21 (50.0)	
	Ukraine	1 (2.4)	
	Uzbekistan	3 (7.1)	
	Vietnam	1 (2.4)	
Academic performance	High	30 (71.4)	
	Low	12 (28.6)	
Social status	≤ Middle	30 (71.4)	
	>Middle	12 (28.6)	
Height (cm)			164.4 (9.6)
Weight (kg)			55.3 (12.2)

Box 1. The qualitative evaluation of R-Ma Bot.
Question: "Please feel free to describe what you liked or could improve while using the chatbot R-Ma Bot."

■ What they liced

■ What they liked

➢ Usefulness

- "It was good to know what nutrients were in the beverages I drank."
- "It was good that he analyzed the drinks I drank and mentioned the negatives of beverages accurately and in detail."
- "I looked back on my habit, and I tried to intentionally reduce it."
- "It helped me lose weight."
- "I was sad to see the chatbot go because we set goals together and accomplished them."

➢ Friendliness

- "I thought he would be stiff like a robot or wouldn't understand me, but he was friendly, smart and kind."
- "I became attached to it for 2 weeks. If I had a chance later, I would hope to meet him again."
- "R-Ma Bot was friendly."

➢ Ease of use

- "I could easily see how many drinks I consumed per day."
- "It was easy to use."
- "It presented information in a way that was easy to understand."

■ Recommendations for improvement

➢ Unnatural conversation

- "There were many topics and questions that it didn't understand."
- "The talk was too long, so I wished the chat could be shortened."
- "I thought it would feel more human if the tone of speech changed every day."

➢ Need more information
- "I wished there had been a greater variety of beverage choices when finding out the ingredients of beverages."

R-Ma Bot: Read and Manage your health roBot.

Usability: Recruitment, retention, participation rate, and engagement

Recruitment was completed in 2 and 4 days for native and racial and ethnic adolescents, respectively. The target number of 42 participants was recruited, with no dropouts during the screening process. The retention rate was 95.2%, with a total of 40 participants remaining at the end of intervention.

Daily participation rates ranged from 83.3% to 100%, and 26 participants had 100% participation over the 14 days. Ethnicity was the primary characteristic affecting 100% of the participation rate. Participation rates in the native group ranged from 90.5% ($n = 19$) to 100% ($n = 21$), and from 76.2% ($n = 16$) to 100% ($n = 21$) in the racial and ethnic group. Of the participants with 100% participation ($n = 26$), only 30.8% ($n = 8$) were racial and ethnic. Additionally, the average number of missed days for native adolescents was 0.14 ($SD = 0.36$) compared to 1.48 ($SD = 1.86$) days for racial and ethnic adolescents. The lowest participation was on Day 8 of the intervention.

A total of 554 data points were collected from 42 people over 14 days. Based on the 20:00 reminder alarm, 132 (22.5%) were classified as active engagement, with data entered before the reminder, and 422 (71.8%) were classified as passive engagement, with data entered after the reminder.

Effectiveness: Knowledge and behavior change

Before the intervention, 64.3% ($n = 27$) of the 42 participants were aware of nutrition labels, which increased to 92.9% ($n = 39$). A total of 42.9% ($n = 18$) of participants had not previously read nutrition labels, which decreased

Table 2. Scores for each chatbot usability questionnaire item.

Chatbot usability questionnaire items	Range	*M* (*SD*)
The chatbot's personality was realistic and engaging	1.0–5.0	3.6 (0.9)
The chatbot seemed too robotic	1.0–5.0	2.7 (0.9)
The chatbot was welcoming during initial setup	3.0–5.0	4.1 (0.7)
The chatbot seemed very unfriendly	1.0–3.0	1.8 (0.7)
The chatbot explained its scope and purpose well	2.0–5.0	4.1 (0.8)
The chatbot gave no indication as to its purpose	1.0–4.0	1.7 (0.7)
The chatbot was easy to navigate	2.0–5.0	4.0 (0.8)
It would be easy to get confused when using the chatbot	1.0–4.0	2.0 (0.8)
The chatbot understood me well	2.0–5.0	3.7 (0.9)
The chatbot failed to recognize a lot of my inputs	1.0–4.0	1.8 (0.8)
Chatbot responses were useful, appropriate, and informative	2.0–5.0	3.9 (0.7)
Chatbot responses were irrelevant	1.0–3.0	1.7 (0.7)
The chatbot coped well with any errors or mistakes	1.0–5.0	3.4 (1.1)
The chatbot seemed unable to handle any errors	1.0–4.0	2.0 (0.9)
The chatbot was very easy to use	3.0–5.0	4.2 (0.8)
The chatbot was very complex	1.0–3.0	1.5 (0.7)

†Even numbered items (negative items) were reverse-coded according to the formula of the original instruments during the analysis.

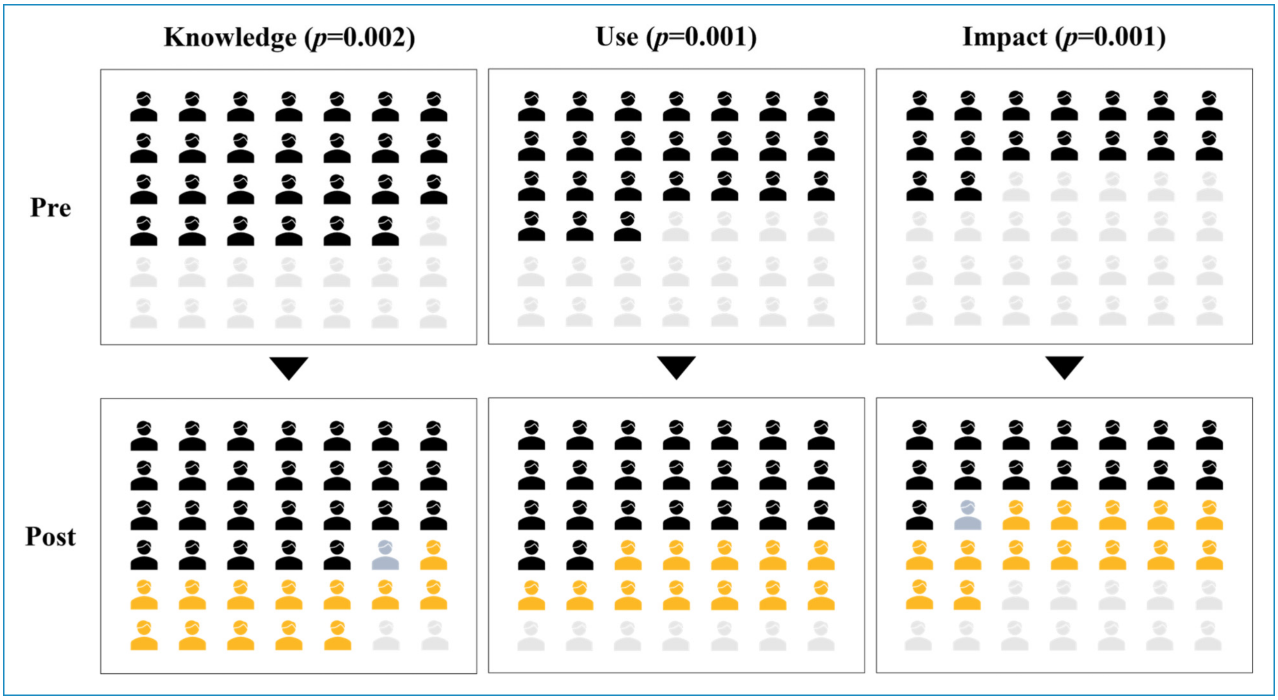

Figure 2. Changes in knowledge, use, and impact of nutrition labels.
†Based on behavioral changes through the intervention, black indicates Yes → Yes, blue indicates Yes → No, yellow indicates No → Yes, and gray indicates No → No. ‡Each statistical significance was confirmed using the binominal McNemar's test.

to 16.7% ($n = 7$) after the intervention. The percentage of participants who stated that nutrition labels influenced their choices increased from 38.1% ($n = 16$) to 69.0% ($n = 29$; Figure 2). The McNemar test for paired binominal data revealed statistically significant changes in all areas of knowledge ($p = .002$), use ($p = .001$), and impact ($p = .001$) before and after the intervention.

During the intervention, participants experienced a significant reduction in weekly average consumption of sugar from beverages, decreasing by 60% from 13.1 ± 20.1 mg to 7.9 ± 12.8 mg ($t = 2.26$, $p = .029$) (Table 3). However, sodium and caffeine intake did not decrease. Participants with a 100% participation rate displayed significantly reduced sugar intake ($t = 2.24$, $p = .034$), whereas the sugar intake of those with a lower participation rate did not decrease. For further analysis, we compared the intake between the native and racial and ethnic groups, controlling for 100% participation, and found that only racial and ethnic adolescents significantly reduced their sugar ($t = 2.55$, $p = .019$) and caffeine ($t = 2.56$, $p = .019$) intake in Week 2 compared to Week 1. However, the average intake of sugar, sodium, and caffeine for racial and ethnic adolescents was approximately twice that of natives in Week 1.

Discussion

The R-Ma Bot displayed high usability, with the high usability scores and also demonstrated high usefulness, friendliness, and ease of use which are the themes that emerged from the qualitative data. These results were supported by the feasibility of the chatbot, such as its rapid recruitment, high retention, and high participation rates. The participants also showed a significant reduction in sugar intake from carbonated and/or energy drinks and an increase in the knowledge, use, and influence of nutrition labels while choosing beverages, confirming the potential effect of the R-Ma Bot for health promotion.

These results may be due to the application of BCT strategies based on behavior change theory in the chatbot. Most prior chatbot studies have not been based on a theoretical framework related to behavior change.[12] In a study examining attitudinal, intentional, and behavioral changes regarding meat consumption among college students, the effectiveness of chatbots applying BCTs was confirmed, although emotional support showing greater effect than information provision.[16] However, rather than applying a single BCT, it is recommended that a comprehensive integration of BCTs into the design of chatbot interventions to enhance effective behavioral changes.[14] Similar to our study, a research demonstrating effectiveness in dietary changes and weight loss among adults has integrated capability-enhancing strategies with other BCTs.[15] While they conceptualized capability as practical skills, unlike our study, the mechanism that addresses the fundamental elements of behavior change is similar. Despite being a short 2-week pilot study, it can be inferred that the behavior

Table 3. Weekly average consumption of sugar, sodium, and caffeine from beverages.

	n	Sugar (mg)				Sodium (mg)				Caffeine (mg)			
		Week 1 *M (SD)*	Week 2 *M (SD)*	*t*	*p*	Week 1 *M (SD)*	Week 2 *M (SD)*	*t*	*p*	Week 1 *M (SD)*	Week 2 *M (SD)*	*t*	*p*
Total	42	13.1 (20.1)	7.9 (12.8)	2.26	.029*	10.1 (16.5)	8.6 (24.3)	0.34	.739	8.6 (10.8)	6.0 (12.3)	1.06	.296
100% Participation	26	11.8 (18.7)	7.4 (11.5)	2.24	.034*	8.1 (14.1)	5.2 (13.1)	1.02	.318	7.5 (10.0)	5.3 (8.7)	1.16	.259
<100% Participation	16	15.3 (22.7)	8.6 (15.0)	1.26	.227	13.3 (20.0)	14.1 (35.8)	−0.07	.946	10.4 (12.1)	7.2 (16.9)	0.56	.586
Native group	21	9.4 (18.1)	8.2 (12.7)	0.46	.652	6.4 (11.0)	12.5 (32.6)	−0.84	.411	5.5 (6.5)	8.5 (16.3)	−0.89	.385
Racial and ethnic group	21	16.8 (21.8)	7.6 (13.2)	2.55	.019*	13.8 (20.3)	4.7 (10.8)	1.86	.078	11.7 (13.3)	3.5 (5.7)	2.56	.019*

†Paired *t*-test was performed to compare the average intake between Week 1 and Week 2.

$^*p < .05$.

change observed in our study was likely due to the provision of customized information based on individual intake. This provides scientific evidence that integrating various BCTs into chatbots is also effective for adolescents.

The R-Ma Bot was considered friendly by participants because of its user-friendly design, adolescent speech patterns, emoticons, and broad conversation categories, including adolescent interests. Chatbots have the advantage of being conversational; therefore, their usability depends on whether sufficient interaction has occurred.[29] Unlike existing chatbots that cover limited topics or have chatbot-centered conversations,[16,30] implementing the R-Ma Bot was unconstrained, allowing access to adolescents' interests beyond just beverages at any time. This approach facilitated the initiation of conversations and likely contributed to the positive evaluation received. Employing reminders, a commonly utilized functional element, proved effective in enhancing both retention and participation rates. Given that only a small percentage (22.5%) of the data was categorized as active engagement, the utilization of reminders might have played a role in stimulating participation, even if the nature of engagement remained passive.

Limitations

Some participants' feedback suggested that the conversations were unnatural and required improvement. Minimizing language constraints and enabling logical communication is necessary for increasing the usability of chatbots.[31] In this study, the free platform was utilized for the cost-effective accessibility of the chatbot, which limited the implementation of high-level artificial intelligence chatbot technology. Technological improvements are needed to ameliorate language use and communication logic limitations, suggesting the need to promote interdisciplinary research involving fields such as nursing and engineering. Such feedback highlights the importance of qualitative data, which cannot be fully captured by quantitative data alone. However, as seen in the non-response cases in this study, optional open-ended questions can introduce response bias.[32] To mitigate this risk, it is recommended to avoid the use of broad questions or large response boxes that may increase participant burden[32] and to opt for data collection at intervals rather than only at the final stage of the study.[33]

Pilot studies are considered a key stage in the development and evaluation of population health interventions. However, there remains a risk of generalizability biases when interpreting the results of such studies.[34] In this context, the findings of our study should be interpreted with caution. Specifically, this study was conducted with a single group, pre- and posttest design and the analysis was based on self-reported intake data, which result in the lack of ability to draw causal inferences about the effectiveness of interventions as well as measurement errors from the actual intake. Future studies should incorporate the improvements identified in this pilot study and verity the intervention's effectiveness through a randomized controlled trial (RCT). Additionally, to ensure the validity of the intervention effects, employing objective beverage intake measurement methods is also suggested.[35]

Unlike the high retention observed in this study, the low participation rates of the racial and ethnic group are important. Racial and ethnic adolescents consumed more carbonated and/or energy drinks, indicating health vulnerabilities, and the significant behavioral change after the intervention, despite low participation rates, supports the need for health promotion programs for them. However, this study has the limitation of not identifying the reasons for the low participation rates among some participants during the intervention. Although this was not evident in our results, previous research has shown that sociocultural factors, such as income, region of residence, and race and ethnicity influence digital disparities, including digital access.[36,37] Therefore, future research should conduct evaluations that includes these considerations and explore strategies to maintain high participation rates for sustainability.

Conclusions

This study aimed to incorporate BCTs based on behavior change theory into a chatbot to reduce the consumption of carbonated and/or energy drinks among adolescents. The R-Ma Bot showed good feasibility with high retention and participation rates during the intervention and was highly useful for promoting healthier beverage intake. It is particularly important to incorporate strategies that enhance adolescents' capabilities by providing knowledge and skills. Additionally, to prevent dropout from digital interventions, youth-friendly designs and engaging approaches are essential. Further studies are suggested to improve the intervention through a comprehensive understanding of both quantitative and qualitative data. Moreover, it is recommended to conduct RCTs to evaluate the effectiveness of this intervention.

Acknowledgments: The authors would like to thank SH Hong and HJ Yim for their assistance with the chatbot development. The first author Lee received a scholarship from Brain Korea 21 FOUR Project funded by National Research Foundation (NRF) of Korea, Yonsei University College of Nursing.

Author's note: Hyeyeon Lee is also affiliated with the College of Nursing, Kosin University, Busan, South Korea.

Contributorship: Jisu Lee: study conception and design, data collection, data analysis and interpretation, drafting of the article; Hyeonkyeong Lee: study conception and design, data analysis and interpretation, and critical revision of the article; Hyeyeon Lee: study conception and design, data interpretation, and critical revision of the article.

Data availability: The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Declaration of conflicting interests: The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Ethical approval: Ethics approval was obtained from the Yonsei Institutional Review Board (IRB-4-2022-1355) and registered in the cris.nih.go.kr (KCT0008114).

Funding: The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This research was supported by the Basic Science Research Program through the National Research Foundation of Korea (NRF), funded by the Ministry of Education under Grant No. NRF-2020R11A2069894).

Guarantor: HKL.

ORCID iD: Hyeonkyeong Lee https://orcid.org/0000-0001-9558-7737

Supplemental material: Supplemental material for this article is available online.

References

1. Korea Disease Control and prevention Agency (KDCA). Korean youth health behavior survey. 2021.
2. Hu H, Song J, MacGregor GA, et al. Consumption of soft drinks and overweight and obesity among adolescents in 107 countries and regions. *JAMA Network Open* 2023; 6: e2325158–e2325158.
3. Sylvetsky AC, Visek AJ, Halberg S, et al. Beyond taste and easy access: physical, cognitive, interpersonal, and emotional reasons for sugary drink consumption among children and adolescents. *Appetite* 2020; 155: 104826.
4. Wikoff D, Welsh BT, Henderson R, et al. Systematic review of the potential adverse effects of caffeine consumption in healthy adults, pregnant women, adolescents, and children. *Food Chem Toxicol* 2017; 109: 585–648.
5. Russo RG, Northridge ME, Wu B, et al. Characterizing sugar-sweetened beverage consumption for US children and adolescents by race/ethnicity. *J Racial Ethn Health Disparities* 2020; 7: 1100–1116.
6. Schneider S, Schilling L and Osenbrügge N. Determinants of soft drink consumption among children and adolescents in developed countries – a systematic review. *Cent Eur J Public Health* 2021; 29: 290–300.
7. Haidar A, Carey FR, Ranjit N, et al. Self-reported use of nutrition labels to make food choices is associated with healthier

dietary behaviours in adolescents. *Public Health Nutr* 2017; 20: 2329–2339.
8. Kim Y, Lee H, Lee H, et al. Social determinants of health of multicultural adolescents in South Korea: an integrated literature review (2018~2020). *Research in Community and Public Health Nursing* 2021; 32: 430–444.
9. Haddock A, Ward N, Yu R, et al. Positive effects of digital technology use by adolescents: a scoping review of the literature. *Int J Environ Res Public Health* 2022; 19: 14009.
10. Dienlin T and Johannes N. The impact of digital technology use on adolescent well-being. *Dialogues Clin Neurosci* 2020; 22: 135–142.
11. Zhang J, Oh YJ, Lange P, et al. Artificial intelligence chatbot behavior change model for designing artificial intelligence chatbots to promote physical activity and a healthy diet: viewpoint. *J Med Internet Res* 2020; 22: e22845.
12. Oh YJ, Zhang J, Fang M-L, et al. A systematic review of artificial intelligence chatbots for promoting physical activity, healthy diet, and weight loss. *International Journal of Behavioral Nutrition and Physical Activity* 2021; 18: 160.
13. Michie S, van Stralen MM and West R. The behaviour change wheel: a new method for characterising and designing behaviour change interventions. *Implement Sci* 2011; 6: 42.
14. Michie S, Richardson M, Johnston M, et al. The behavior change technique taxonomy (v1) of 93 hierarchically clustered techniques: building an international consensus for the reporting of behavior change interventions. *Ann Behav Med* 2013; 46: 81–95.
15. Maher CA, Davis CR, Curtis RG, et al. A physical activity and diet program delivered by artificially intelligent virtual health coach: proof-of-concept study. *JMIR Mhealth Uhealth* 2020; 8: e17558. 20200710.
16. Carfora V, Bertolotti M and Catellani P. Informational and emotional daily messages to reduce red and processed meat consumption. *Appetite* 2019; 141: 104331. 20190702.
17. Shagiwal SS, Groenestein E, Schop-Etman A, et al. Effectiveness of behavioral interventions and behavior change techniques for reducing soft drink intake in disadvantaged adolescents: a systematic review and meta-analysis. *Obesity Science & Practice* 2020; 6: 708–734.
18. Ezendam NP, Brug J and Oenema A. Evaluation of the web-based computer-tailored FATaintPHAT intervention to promote energy balance among adolescents: results from a school cluster randomized trial. *Arch Pediatr Adolesc Med* 2012; 166: 248–255. 20111107.
19. International Organization for Standardization. ISO ergonomics of human-system interaction – Part 210: Human-centred design for interactive systems. 2010.
20. Creswell JW, Clark VLP. Core mixed methods designs. In: Salmon H, Scappini J, DeRosa K, et al. (eds) *Designing and conducting mixed methods research*, 3rd ed. USA: Sage Publications, 2017, pp.51–100.
21. Kang YK, Jang YN, Bae SY, et al. A meta-analysis of the effects of artificial intelligence education programs on adolescents. *Journal of Korean Association for Educational Information and Media* 2021; 27: 1273–1294.
22. Gabrielli S, Rizzi S, Bassi G, et al. Engagement and effectiveness of a healthy-coping intervention via chatbot for university students during the COVID-19 pandemic: mixed methods proof-of-concept study. *JMIR Mhealth Uhealth* 2021; 9: e27965. 20210528.
23. Cimpeanu I-A. Business activities using chatbots. *Informatica Economica* 2021; 25: 38–51.
24. Holmes S, Moorhead A, Bond R, et al. Usability testing of a healthcare chatbot: can we use conventional methods to assess conversational user interfaces? In: Proceedings of the 31st European conference on cognitive ergonomics. BELFAST, UK: Association for Computing Machinery, 2019, pp.207–214.
25. Holmes S, Bond R, Moorhead A, et al. Towards validating a chatbot usability scale. In: Marcus A, Rosenzweig E and Soares M (eds) *Design, User Experience, and Usability. HCII 2023. Lecture Notes in Computer Science*. Cham: Springer, 2023, pp.321–339.
26. KDCA. Korea national health and nutrition examination survey. 2021.
27. Caffeineinformer. Caffeine content of drinks, https://www.caffeineinformer.com/the-caffeine-database (2023, accessed 19 July 2023).
28. Goodman E, Adler NE, Kawachi I, et al. *MacArthur scale of subjective social status–youth version*. APA PsycTests, 2001.
29. Milne-Ives M, Cock C, Lim E, et al. The effectiveness of artificial intelligence conversational agents in health care: systematic review. *J Med Internet Res* 2020; 22: e20346. 20201022.
30. Fadhil A, Wang Y and Reitere H. Assistive conversational agent for health coaching: a validation study. *Methods Inf Med* 2019; 58: 9–23. 20190522.
31. Fadhil A. Can a chatbot determine my diet?: Addressing challenges of chatbot application for meal recommendation. 2017.
32. Zuell C, Menold N and Körber S. The influence of the answer box size on item nonresponse to open-ended questions in a web survey. *Soc Sci Comput Rev* 2014; 33: 115–122.
33. Miller A and Dumford A. Open-ended survey questions: item nonresponse nightmare or qualitative data dream? *Surv Pract* 2014; 7: 1–11.
34. Thabane L, Cambon L, Potvin L, et al. Population health intervention research: what is the place for pilot studies? *Trials* 2019; 20: 309.
35. Aaby D and Siddique J. Effects of differential measurement error in self-reported diet in longitudinal lifestyle intervention studies. *International Journal of Behavioral Nutrition and Physical Activity* 2021; 18: 125.
36. McBain RK, Cantor JH, Kofner A, et al. Ongoing disparities in digital and in-person access to child psychiatric services in the United States. *J Am Acad Child Adolesc Psychiatry* 2022; 61: 926–933.
37. Ferretti A, Hubbs S and Vayena E. Global youth perspectives on digital health promotion: a scoping review. *BMC Digital Health* 2023; 1: 25.