Chatbot Design for a Healthy Life to Celiac Patients: A Study According to a New Behavior Change Model

Eythar Alghamdi[1], Reem Alnanih[2]
Computer Sciences Department, Faculty of Computing and Information Technology[1, 2]
King Abdulaziz University, Jeddah, Saudi Arabia[1, 2]
University of Bisha, Bisha, Saudi Arabia[1]

***Abstract*—There is an absolute need for technology in our daily life that makes people busy with their smartphones all day long. In the healthcare field, mobile apps have been widely used for the treatment of many diseases. Most of these apps were designed without considering health behavior change models. Celiac disease is a significant public health problem worldwide. In Saudi Arabia, the incidence of celiac disease is 1.5%. Celiac patients have a natural demand for resources to facilitate care and research; however, they have not received much attention in the field of healthcare apps. This study introduced a new health behavior change model based on the existing common models and adapted it to the use of technology for the changing behavior of celiac patients towards healthy suitable food. As proof of concept, the new model was applied to the WhatsApp chatbot for patients with celiac disease. To test the impact of the chatbot, 60 Saudi celiac patients participated in three steps. First, they completed a pre-test questionnaire. Then, the participants were divided into two groups: the control group, which was left without any intervention, and the test group, who used the chatbot for 90 days. Finally, all participants completed the post-test questionnaire. The results confirmed a significant statistical difference between both groups, and the test group improved their healthy life in terms of eating habits, reduced celiac symptoms, and commitment to the treatment plan.**

***Keywords*—Celiac disease; health behavior changes models; healthcare apps; user-centered design; experiment test; WhatsApp chatbot**

I. Introduction

There is an absolute need for technology in daily life. Everyone is dependent on technology to the point that we cannot stay without it; starting from communicating with others and ending monitoring of health. One of the most popular technologies worldwide is mobile, and the number of people around the world who use mobile in 2021 is 5.27 billion users, accounting for 66.85% of the total global population, every one of them spends an average of 145 minutes daily online [1][2]. In addition, statistics show that mobile users in the United State rely on their mobile devices to make decisions about many things, such as making decisions about purchasing a product, choosing a place to live, or choosing daily activities [3].

In the healthcare field, mobile apps have been used widely for many diseases, such as atrial fibrillation and diabetes. It has also been used to control or change unhealthy behaviors, such as smoking and alcohol drinking [4-7]. Most healthcare apps aim to monitor patients' behaviors and follow their improvement [8]. On one hand, most of these apps were designed without considering the health behavior change models such as the health belief model (HBM), theory of planned behavior (TPB), diffusion of innovation theory (DOI), social norms theory (SNT), and transtheoretical model (TTM) [8]. This cast doubts their effectiveness in controlling user behavior or changing it [9]. On the other hand, a few apps are based on one of the health behavior change models, but some studies have proven that they are ineffective in avoiding relapse and because there are many flaws in these models [9].

Health apps depend on the user to enter most of the data, and if the user does not commit to using them daily by entering the required data, it may become useless. A study has proven that the more apps on the mobile device, the less the user uses them [10]. This means that if a mobile user has 20 apps on his mobile device, he uses approximately 8-10 of them actually. In addition, same statistics showed that mobile users use social networks more than other apps, spend half of the mobile use time on it, and they use it daily and for extended hours.

A social network service is an online platform that people use to build social relations with other people by sharing their interests. WhatsApp is the most used social network globally, with 1.5 billion monthly active users [10]. WhatsApp allows the user to communicate with other users using writing messages, voice notes, voice calls, and video calls through simple encrypted interfaces. WhatsApp also supports the conversational system, which is a computer system intended to converse with a human in a manner similar to a human using writing, speaking, or gesturing. The conversational system actions are based on user inputs using both short-and long-term user knowledge [11]. The chatbot is a conversational system supported by WhatsApp. Chatbots have been widely used in customer service. In the field of health, chatbots are used on a small scale to spread health information, deny rumors, and correct false information [12, 13].

Celiac disease is a significant public health problem worldwide. It is an autoimmune enteropathy resulting from the interplay between environmental and genetic factors, which is affected by eating gluten-containing grains [14]. The current worldwide prevalence of celiac disease is between 0.7% and

1.4%, and it affects women more than men and children more than adults [15]. In Saudi Arabia, the incidence of celiac disease in adults and children is 1.5%, which is at least twice the average prevalence rate in Europe and North America [16, 17]. Although it is a high rate, there is a study that confirmed that the incidence of celiac disease in Saudi Arabia is higher than 1.5% and reaches 2.7% [18]. However, celiac patients have not received much attention in the field of healthcare apps; rather, they rely heavily on nutrition apps, and only 5% of specialists use social networks, websites, and nutrition apps to educate celiac patients [19, 20]. In contrast, celiac patients have a natural demand for web-based resources to facilitate care and research, especially on mobile devices [21].

Therefore, this paper aims to take the benefit of user experience toward using social media but, instead of taking this user experience randomly, we take it based on a scientific model. All this is motivated by the human sense to improve the life quality of celiac patients since the potential benefits of the proposed solution are painless and easy life for them. This paper focuses on the problem from two perspectives, as follows:

- The user's perspective: The user uses many apps in order to change his/her health behavior, but a large number of apps on the mobile cause him to become distracted, which makes him gradually tend to quit using those apps. In addition, healthcare apps built based on one of the common health behaviors change models are negatively affected by their limitations. Therefore, it would be better to design a model that avoids the limitations of common models and benefits from the apps that the user already uses to change behavior.
- Developer's perspective: The developer wants to develop an app that achieves widespread and supports the sustainability of continuous use by users, especially if its target changes behavior. It has been noticed that social networks are the most used mobile apps. Therefore, it would be better to use social networks to change healthy behaviors rather than developing a new mobile app.

In this paper, the authors proposed a new health behavior change model that can be adapted to the technology. The proposed model is based on common health behavior change models in terms of the defined criteria of objectives for using each one, method of implementation, duration of implementation, advantages, and disadvantages. As a proof of concept, the proposed model was applied to the chatbot through WhatsApp. The chatbot is designed for celiac patients as a scope of the research, and gamification is used as an engagement boost tool. This study's main contributions are the new health behavior change model and the benefit of social media in behavior change. The implications of these contributions are to the creation of a health behavior change model that avoids the shortcomings of existing models. Also, taking advantage of the long hours that users spend using social media.

The proposed design was empirically tested with 60 celiac patients divided into two groups for three months (90 days) based on a study that proved that celiac patients' condition could improve from one month to six months [22, 23].

The rest of the paper is organized as follows: The second section contains the literature review, and the third section provides the data collection and analysis. The fourth section explains the proposed model, and the fifth section shows the prototyping and design phases. The sixth section presents the details of the experimental design, the seventh section discusses the results, and the eighth section concludes the paper.

II. Literature Review

This section divides the literature review into several subsections in terms of the health behaviors change models, their applications in the healthcare field, mobile health apps, WhatsApp and chatbot, and the methods used to collect and analyze the data.

A. Health Behavior Change Models

Social and behavioral sciences researchers have invented different models to change people's behavior in the health field. The most common models are the HBM, TPB, DOI, TTM, and SNT [24-28].

Although these models have some advantages, they have many disadvantages: for example, HBM works only with a fatal disease [24], TPB is based on motivation and ability [25], DOI works better with the adoption of behaviors rather than cessation or prevention of behaviors [26], SNT focuses only on social influences and their impact on an individual's behavior [28], while TTM operates on the assumption that taking years to change behaviors [27]. In addition, many of these models do not consider the maintenance of behavior, but rather focus on initiating behavior.

B. Health Behavior Change Models Applications

To the best of our knowledge, no studies have explored the impact of the application of health behavior change models on celiac patients. One study [29] found that TPB was used to measure celiac patients' beliefs about a gluten-free diet. However, it did not measure the effect of TPB on patients with celiac disease in terms of health behavior change.

The authors in [30-33] used HBM in their studies. The authors in [30] targeted the asymptomatic hyperuricemia patients, but measured 5 of HBM constructs only, the follow-up was short, and participants were from one community. The authors in [31] targeted the obesity patients, but the sample was small because of the low response rate and data collection was from a single campus. The authors in [32] targeted the gestational diabetes patients, but there was a lack of related works on which the study relied, and they used self-reporting. The authors in [33] targeted the smokers, but gathering information was by self-reporting, the study relied on people's ability to remember past behaviors accurately, which cannot be trusted, and due to some legal, ethical, and social reasons in the study community, it is not possible to trust that the answers obtained were accurate.

The authors in [34-36] used TPB in their studies. The authors in [34] targeted the smokers, but the implementation

was in the tobacco belt, which may affect the reactions of the participants, and the study measured the smoking behaviors before the campaign only, it did not measure it after the campaign. The authors in [35] targeted the type two diabetes patients, but the study did not take the patient's diversity into account and the social desirability bias might affect the accuracy of the data. The authors in [36] targeted the cervical cancer patients, but the study did not employ an experimental design to establish the causal effect relations among variables.

The authors in [37-38] used DOI in their studies. The authors in [37] targeted the smokers, but the implementation relied mainly on community awareness in a community that accepts smoking and there was no monitor of participants to ensure adherence. Authors in [38] targeted the maternal and neonatal health, but the health systems problems, Infrastructure problems, and cultural constraints affected the study result.

The authors in [39-41] used TTM in their studies. The authors in [39] targeted the smokers, but the study did not used the blood tests to measure nicotine before and after the experiment, the study did not include all TTM stages, the sample was small, and the follow-up was short. Authors in [40] targeted the cardiovascular patients, but the intervention was based on individuals' readiness for behavioral changes, and it had not part in making them ready, and the post-test was not applied. Authors in [41] targeted the obesity patients, but gender was not considered as a mediating variable, and there was the sample loss during the study.

Finally, the authors in [42-44] used SNT in their studies. The authors in [42] targeted the smokers, but the study linked social norms and the use of tobacco products with a temporal relationship, which in turn changes and therefore it is not possible to establish the relationships between them. Authors in [43] targeted the type two diabetes patients, but there was a lack of real behavioral measurements, which limits the generalization of these results. Authors in [44] targeted the people who eat junk food, but the follow-up was very short.

From our review and analysis of the methodologies that were used in previous studies, the aim of each model and the most appropriate case for its use became clear. Also, the shortcomings of each model were discovered. Table I describes the proposed new model compared to the other common models in terms of a set of criteria such as the number of processes, the healthcare domain, problem-solving issues, applying the technology, using social networks, taking advantage of social influences, and the ability of the model to work with behavior adoption, behavior prevention, and behavior maintenance.

C. Healthcare Mobile Apps

Based on the studies published in the field of health apps, it was noticed that many of those apps work based on one or more of the health behavior change models in order to track behavior [8]. The authors in [4] used SNT in their app and targeted the atrial Fibrillation disease patients, but a very small sample used the app (only 2 patients). Authors in [5] used TTM in their app and targeted the Type two diabetes patients, but it was limited to patients with high motivation to use a mobile app for self-management. Authors in [7] used SNT in their app and targeted the alcohol drinkers, but the sample was from rural and urban without consider the digital divide between them. Authors in [45] used TPB in their app and targeted the hospitalized smokers, but the sample size was relatively small, and the follow-up was short. Authors in [46] used DOI in their app and targeted the gestational diabetes mellitus patients, but the number of interviews was limited. Authors in [47] used TTM in their app and targeted the children with diabetes, but the app did not include all TTM stages.

TABLE I. THE COMPARISON CRITERIA BETWEEN MODELS

Model	Process	Problem	Healthcare Domain	Technology	Social Networks	Social Influences	Behavior Adoption	Behavior Prevention	Behavior Maintenance
HBM	Six constructs	Diabetes, Hyperuricemia, Obesity Smoking.	√	√	√	X	√	√	X
TPB	Six constructs	Cancer, Diabetes, Smoking.	√	X	X	X	√	X	X
DOI	Five categories	Maternal & Neonatal health, Smoking,	√	X	X	X	√	X	X
TTM	Six stages	Cardiovascular, Obesity, Smoking.	√	X	X	X	√	√	√
SNT	Six phases	Diabetes, junk food eating, Smoking.	√	X	X	√	√	X	X
The new model	Four stages	Celiac, Chronic disease.	√	√	√	√	√	√	√

D. WhatsApp and Chatbot

WhatsApp is the most widely used social and communication app globally. User engagement owing to its features, such as free calls and messages, the user interfaces are simple, easy to use, and messaging encryption [10]. The chatbot is a conversational agent that interacts with users using natural language [48]. Chatbot inhabits platforms such as Facebook, Messenger, WhatsApp, Telegram, iMessage, and even websites. Users communicate with a chatbot via the chat interface, such as talking to a real person, chatbot interprets, and processes the user's words and provides an instant answer.

Different chatbots have been developed using text communication, and they have been used in various domains such as customer service, education, and website help [48]. Chatbots have been used in the healthcare field for many reasons, like:

- Change behavior as in food consumption chatbot [11] on Facebook Messenger and smoking cessation chatbot [49] on Telegram.
- Sharing health information as in COVID-19 chatbot [12] on WhatsApp and Saudi health ministry chatbot [13] on Telegram.
- Monitoring Patients as in chronic conditions' teenagers chatbot [50] by Text Message.

As a summary for this section, the common existing models for health behavior change have critical limitations that effect of their applications negatively. Therefore, this paper intends build a new model to avoid those limitations. Also, it's notice that the use of social media to change health behavior is an idea that was previously applied, but not applied to change the behavior of patients with chronic diseases. Therefore, this paper intends to take advantages and use social media to change behavior of chronic diseases patients.

III. Data Collections and Analysis

Three different collection methods were used in this study: 1) Interview with celiac expert users. 2) Questionnaire distributed to patients with celiac disease. 3) Studying exciting changing behavior models in the literature review. The reason is to collect the information data from different angles and draw a complete picture of the existing problem. The following subsections describe each method and their results.

A. Existing Health Behavior Models

Five common existing health behavior models were explored, examined, and analyzed. The existing health behavior models were HBM, TPB, TTM, DOI, and SNT. The goal was to determine the similarity, differences, necessary stages, and unnecessary stages to build a new health behavior change model that is compatible with the needs of patients with celiac disease. To analyze the five-existing model, three steps were followed:

1) Finding similarities between the three models (HBM, TPB, TTM).

2) Dispense with unnecessary stages in the three models (HBM, TPB, and TTM).

3) SNT and DOI were added to the three models (HBM, TPB, and TTM).

B. Interviews

Interviews were conducted with celiac experts to find out the following:

- The most common celiac symptoms.
- The most common treatment plan for celiac.
- The difficulties that celiac patients face in treatment plan commitment.

The celiac experts were divided into four categories as follows:

- Group 1: A patient diagnose with celiac four years ago or more.
- Group 2: A patient's parent who was diagnosed with celiac four years ago or more.
- Group 3: A dietitian who has supervised the diet of celiac patients for four years or more.
- Group 4: A gastroenterologist who has treated celiac patients for four years or more.

The reason of chosen four years or more as a criterion are:

1) The period required to change behavior.

2) The period required for a person to become an expert in a field.

It has been proven that any behavior can be changed within 18 days and not more than 254 days (less than a year) [51, 52]. It has also been proven that a person can become an expert in a field by practicing for not less than 10,000 hours [53], which can be considered as 3.9 years if the practice is daily, as in the case of chronic diseases, for 7 hours.

Ten questions were designed and prepared for the interviews. Before the interviews were conducted, a pilot test was conducted with an expert to test the suitability of the questions. The expert is a female celiac patient since 2015 and lives in Madinah, Saudi Arabia. She helps new patients adhere to their treatment plan by providing educational lectures about celiac and how to cope with it. After expert-approval, the interviews were conducted by phone, due to the social distancing measures that were imposed in the summer of 2020– the time of the interviews conducted- due to the spread of COVID-19. The interviews were conducted with four celiac patients, two celiac patients' parents, two dietitians, and three gastroenterology consultants. 54.5% of the interviewees were females, 45.5% were males. 36.4% of the interviewees had Ph.D. degree, 27.3% had bachelor's degree, 18.2% had master's degree, 9.1% had degree less than a bachelor's degree, and 9.1% prefer not to answer the question. 100% of the interviewees believed that the best treatment way for celiac is Gluten-free diet. The celiac symptoms -based on the interviewees' opinions- were stomachache, bloating, diarrhea, nausea, headache, and fatigue. 100% of the interviewees confirmed that the patients never experience the same symptoms after commit the treatment way. Difficulties that the

patients meet -based on the interviewees' opinions- were lack of alternatives, the need to sessions with a dietitian, and lack of support groups.

C. Questionnaire

The questionnaire was distributed among different celiac patients' communities targeting the celiac patients and their relatives to answer the following questions:

- How are patients ranked as the most common physical and psychological symptoms based on the fastest appearance when they eat gluten?
- How are patients ranked as the most common physical and psychological symptoms based on the fastest disappearance when they stop eating gluten?
- What are the patients' preferences regarding the use of technology?

A total of 137 responses were received. The questionnaire was divided into five sections, and each section contained several questions. Before the questionnaire was distributed, an expert tested content validity and face validity. Many comments were received and considered before being distributed.

For reliability, one week after the questionnaire was administered. Sixty-nine participants agreed to retake the questionnaire (50.36% of the total participants). The questions' results were tested by finding the correlation between the results of the first time the questionnaire was submitted and the results of the second time. The overall correlation was 0.9749042904, which indicates a very high positive correlation.

When finding the correlation for each question separately, the results ranged between 0.7967715443 and 1, which made the correlation moderate, and reliability was acceptable. The above results proved the following:

- 83.8% of patients are female.
- 87.6% of patients are Saudi.
- 46% of the patients are 29 years old or less.
- 83.2% of the patients lived in three regions: Riyadh (34.3%), Makkah (29.2%), and Eastern (19.7%) regions.
- 75.18 of the patients lived in six cities: Riyadh (30.7%), Jeddah (22.63%), Madinah (8.03%), Dammam (5.11%), Dhahran (5.11%), and Makkah (3.6%).
- When patients eat gluten, they suffer from stomachache, bloating, diarrhea, nausea, headache, and fatigue.
- When patients suffer from eating gluten symptoms, they feel a desire to isolate themselves, worry that symptoms get worse, fear of what others think about them, that they are a burden to others, shame about their illness, depression, and the desire to cry.
- 67.9% of patients think that having a health application will benefit them.
- All patients can read Arabic.
- 82.5% of patients prefer WhatsApp over any other social media apps.
- Patients want the mobile application to provide alternative products, gluten-free recipes, medical and psychological follow-up, gluten-free diet suggestions, and periodic reports.

IV. Proposed Model

To build a new health behavior change model, the focus has been on, in the first hand, the users and their needs, and, in the second hand, the software development life cycle. Therefore, user-centered design (UCD) has been used as a model that focuses on the users' needs and the waterfall model as a software development life cycle model. In the following subsection, details about the proposed model and its phases, phase duration, and phase quiz.

A. The Chronic Diseases Extended Model

Since the new model extends from five old models, it was decided to name the new model "The Chronic-Disease Extended Model," or "CDEM" in short form. This section provides an overview of CDEM and its phases.

The CDEM involves progress through a series of phases, but progression occurs in a nonlinear fashion. This means that patients can progress from the earlier phase to later ones, regress from the later phase to the earlier one, and recycle through the phases from the last phase to the first one. As shown in Fig. 1, the model consists of the following four phases.

- Phase 1: Preparedness: the phase in which the patient knows what to feel about the disease.
- Phase 2: Readiness: the phase where the patient is ready to deal with the disease.
- Phase 3: Confidence: the phase in which the patient can find solutions for every barrier that faces.
- Phase 4: Continuance: The phase where the patient commits to the treatment plan.

Every phase has its own actions that help reach the phase goal. In addition, a patient is assigned to one of the phases based on his answers on the phase quiz, which is conducted every two weeks.

Fig. 1. Chatbot Processes.

TABLE II. THE PHASES AND THE WAY TO PROGRESS THROUGH THEM

The phase level	Phase character	Progressing status
Preparedness	Person knows what to feel about his disease and be conscious of the need for innovation.	Motivation to start treatment
Readiness	Person feels the seriousness of leaving the disease untreated and knowing treatment ways.	deciding to change behavior
Confidence	Person believes that social norms approve the new behavior and starting the new behavior.	feeling confident in the ability to perform a behavior.
Continuance	Person's perception of the obstacles to performing the behavior and knowing the factors that facilitate the performance of a behavior.	-

Table II presents the phases and ways of progressing through them. The first column in the table shows the phases' level, the second column shows the patient attributes in the different phases, and the last column shows the attribute that makes the patient ready to move to the next phase.

B. Phases Duration

Each phase differs from another in its goals and actions, and each patient differs from the other in his personality and needs. Therefore, it is difficult to predict how long each patient needs at each phase because of the large number of variables. This model suggests that the duration of each phase is a minimum of two weeks and a maximum of three months, based on the duration average of the five models that have been studied. To determine if the patient is ready to move to the next stage, a test called the phase quiz.

C. Phases Quiz

A phase quiz was developed based on the TTM instrument. The TTM instrument was created by Prochaska and DiClemente and validated in different intervention studies, which have focused on multiple health problems. Furthermore, phase quizzes develop based on the short version of the instrument, and its validation has been proven [54]. The reason for this choice is that the short version contains only one question, and the patient never feels bored when he takes the test every two weeks. The phase quiz consists of one question with four clauses that detriment whether patients stay, progress, or regress in the phases based on the sentence the patient's choice. Table III presents the question, its clauses, and action when choosing every clause.

TABLE III. PHASE QUIZ QUESTION AND ITS CLAUSES

Question	Clauses	The action when choosing the clause
Did you do the habit that was recommended to stay away from it in the last two weeks?	Yes, and I feel no bad about it.	The patient is in phase 1: **Preparedness**
	Yes, but I regret it.	The patient is in phase 2: **Readiness**
	No, But I still face some obstacles with no solutions.	The patient is in phase 3: **Confidence**
	No, and I totally committed to the treatment plan.	The patient is in phase 4: **Continuance**

V. PROTOTYPING AND DESIGNING

For proof of concept, the proposed model was applied to the chatbot. This section presents the chatbot analysis and chatbot design.

A. Chatbot Analyzing

To analyze the chatbot, the flowchart diagram as shown in Fig. 2 starts with a welcome message that explains the chatbot's goal and how to interact with it. Then, the pre-processing step, which is a pre-questionnaire to collect demographic information about the participants and their health behaviors and eating habits. After that, the processing includes the phase quiz to determine each participant's phase and interact with them with different actions based on their phase.

For example, in some phases, participants learn about celiac and its complications; in other phases, participants can find a place to order a gluten-free meal. The participants took the quiz phases every two weeks to reassign them to a suitable phase. Finally, after the experiment period was completed, the participants took post-processing, which is a post-questionnaire to collect information about their health behaviors and eating habits.

B. Chatbot Design

This section presents the chatbot prototype, tool used for development, pilot test feedback, and final chatbot design layout.

1) Chatbot prototype: To ensure the correctness of the information provided by the chatbot, information about celiac was collected from the website of the Saudi Ministry of Health. The chatbot prototype was designed using a chatbot design tool called "Botsociety".

2) Deployment service provider: To deploy the chatbot on WhatsApp, based on WhatsApp policy, it is necessary to deal with a licensed company. A search was conducted to identify different licensed companies that work as chatbot deployment service providers. The "Widebot" platform was used for chatbot design after strict comparison between five different companies based on the target customers, supporting WhatsApp, supporting Arabic, having an AI engine, and supporting natural language processing (NLP). For AI, "Widebot" uses its own classifier that combination between exact matching, fuzzy matching, and Naive Bayes classifier. For NLP, "Widebot" uses BERT model.

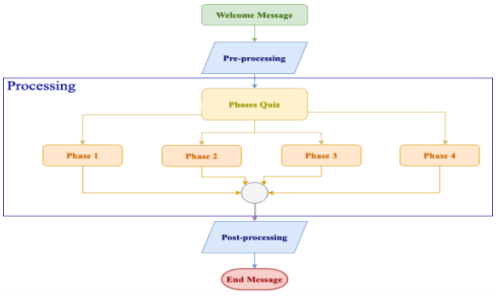

Fig. 2. The Chatbot Flowchart.

VI. Experiment

This section presents the details of the experimental design, including, participants, materials and testing procedure.

A. The Participants

The experiment was conducted on 60 Saudi participants diagnosed with celiac disease. Sixty participants were recruited through online virtual communities of celiac patients in Saudi Arabia. Many patients expressed their desire to participate in the experiment (137 patients). Only 60 participants met the experimental criteria. The criteria were:

- The participants lived in Jeddah. Jeddah city was chosen because it can suggest places that provide gluten-free meals only in Jeddah.
- Participants read and wrote Arabic fluently. Arabic was chosen because Arabic is the language of the chatbot.
- The participant approves to participate in the experiment and commits to it until the experiment ends by signing the consent form.

The 60 celiac patients who participated in the study experiment were split randomly into two groups: the control group and the test group. The control group was left without any intervention, which means that the proposed chatbot was not used, while the test group used the proposed chatbot.

B. The Materials

An empirical study was planned and executed through three steps to determine whether the new model helps patients improve their health and adhere to the treatment plan. In addition, determining whether the idea of benefiting from the user experience of WhatsApp may be beneficial in the health field. The three steps are:

- Step 1: Pre-test questionnaire to collect the participants' data on eating habits, celiac symptoms, and treatment plan commitment.
- Step 2: Formal test to collect the data attributes during the test and record it in the log-in form to evaluate the participants interacting with the chatbot and their time of use.
- Step 3: Post-test questionnaire to collect the participants' data about eating habits, celiac symptoms, and treatment plan commitment to evaluate the participants' experience in using the proposed design as a tool to improve their health.

C. Testing Procedure

The participants in the test group only used the proposed chatbot for three months (90 days) since the improvement in celiac patient health could show in a week to 6 months [22, 23]. The experiment was conducted from December 20, 2020, to March 19, 2021. During this process, the researcher observed the WhatsApp conversations with the participants and collected the data daily. The "Widebot" platform is dedicated for collecting the data. This allows the observer to read the conversations and extract data from them. It also gives him an analysis of the chatbot performance and the users' use of it.

VII. Results and Discussion

This section analyzes the pre-test and post-test questionnaire results.

A. Pre-Test Result

First, the control and test groups answered the pre-test questionnaire on December 19, 2020, and the test group started using the proposed chatbot on December 21, 2020. The pre-test questionnaire consisted of the following four questions:

1) When was the last symptoms date?
2) Do you face difficulties in finding gluten-free products?
3) Do you face difficulties committing to the treatment plan?
4) Did you eat any food that contains gluten in the last two weeks? (Phase quiz)

From the result, it was noticed that 6.67% of the participants in the control group and 6.67% of the participants in the test group felt celiac symptoms last time on the same day they answered the pre-questionnaire. In addition, 30% of the participants in the control group and 30% of the participants in the test group always had difficulty finding gluten-free products. Additionally, 20% of the participants in the control group and 16.67% of the participants in the test group always had difficulties committing to the treatment plan. Finally, 30% of the participants in the control group and 30% of the participants in the test group were in phase one, based on the CDEM.

B. Formal Test

The test group started to use the chatbot only. This step provided a clear picture of the number of chatbot times of use, the chatbot features that user tend to use, and the participants' progress in the CDEM phases. Analyzing the result of this step is useful in improving the chatbot for future research and does not affect the outcome of the experiment.

C. Post-Test Result

After the experiment was completed, both groups of participants answered the post-test questionnaire on March 20, 2021. The post-test consists of the following four questions:

1) When was the last symptoms date?
2) Do you face difficulties in finding gluten-free products?
3) Do you face difficulties committing to the treatment plan?
4) Did you eat any food that contains gluten in the last two weeks? (Phase quiz)

From the result, it was noticed that 0.33% of the participants in the control group felt celiac symptoms last time on the same day they answered the post-questionnaire, but none of the participants in the test group felt celiac symptoms on the same day. In addition, 20% of the participants in the control group always had difficulties finding gluten-free products, but only 0.33% of the participants in the test group always had difficulties finding gluten-free products. Additionally, 6.67% of the participants in the control group

never had difficulties committing to the treatment plan, but 46.67% of the participants in the test group never had difficulties committing to the treatment plan. Finally, 26.67% of the participants in the control group were in phase one based on the CDEM, while none of them were in phase four. However, none of the participants in the test group were in phase one, while 43.33% of them were in phase four.

In general, and based on the previous comparison, it is clear that the status of participants in the control group, who were left without any intervention, did not improve; on the contrary, it regressed in many cases. However, the status of the participants in the test group, who used the proposed chatbot, improved, except concerning difficulty committing to the treatment plan.

D. Discussion

This section discusses the results statistically to reach the results of this study. The first two subsections compare the results of the two groups in the pre-test and post-test by finding the correlation, and ANOVA test.

1) Correlation between the two groups: After obtaining the results, Pearson's correlation analysis between the two groups' results for each question was applied, as it is the most used measure of correlation to show the degree of the linear relationship between variables. Table IV presents the results.

As Table IV shows the correlations between the two groups' answers on the pre-test four questions were large positive correlations between 0.88 and 0.99. The correlations between the two groups' answers to the post-test questions were as follows:

- Medium positive correlation in one out of four questions, equal to 0.76,
- Medium negative correlation in two out four questions between -0.73 and -0.60,
- No correlation in one out of four questions, equal to -0.06.

This means that the two groups were close in their answers in the pre-test. Still, after the test group participants conducted the experiment, the answers differed greatly, which made some correlation results appear negative, which can be said that the intervention applied to the test group (the experiment) affected the results of the group. Therefore, this intervention can help to reach the goal of changing health behaviors using social media.

However, the "no correlation" result for the second question, which was related to facing difficulties in finding gluten-free products, in the post-test was noticed. This result can be attributed to the fact that the 60 participants were reached through some virtual celiac patient communities, in which patients share experiences about gluten-free products, including the free-gluten products' companies, and where those products are available. Therefore, it may be difficult to establish a relationship between the status of the participants in the two groups regarding this question.

2) ANOVA Test between the two groups: An ANOVA test was applied to determine which hypotheses are listed below:

- H0: There is no difference between the behavior of the Celiac patients before and after using the Chatbot based behavior change.
- H1: There is a difference between the behavior of the Celiac patients before and after using the Chatbot based behavior change:

The ANOVA test finds the p-value, a value that determines whether a significant statistical difference exists between the groups.

Table V shows the ANOVA test results of the pre-test for the groups, while Table VI shows the ANOVA test results of the post-test for the groups.

A P-value < 0.05, indicating a significant statistical difference between the groups. In the two tables above, the P-value for the pre-test was 0.77, which is more than 0.05, which means there is no significant statistical difference, but the P-value for the post-test was 0.0, which means there is a statistically significant difference between the groups. Therefore, we can confirm that the proposed chatbot helped celiac patients improve their eating habits.

TABLE IV. THE PRE-TEST AND POST-TEST CORRELATION BETWEEN THE CONTROL GROUP AND TEST GROUP

The list of questions	**Pre-test correlation**	**Post-test correlation**
When was the last symptoms date?	0.96	0.76
Do you face difficulties in finding gluten-free products?	0.88	-0.06
Do you face difficulties committing to the treatment plan?	0.98	-0.73
Did you eat any food that contains gluten in the last two weeks? (Phase quiz)	0.99	-0.60

TABLE V. ANOVA TEST RESULT OF THE PRE-TEST FOR THE CONTROL GROUP AND TEST GROUP

Source of Variation	**Sum of Squares**	**Degrees of Freedom**	**Mean Square**	**F-statistic value**	**P-value**
Between Groups	0.06666666667	1	0.06666666667	0.08618127786	0.7701371105
Within Groups	44.86666667	58	0.7735632184		
Total	44.93333333	59			

TABLE VI. ANOVA TEST RESULT OF THE POST-TEST FOR THE CONTROL GROUP AND TEST GROUP

Source of Variation	**Sum of Squares**	**Degrees of Freedom**	**Mean Square**	**F-statistic value**	**P-value**
Between Groups	28.03333333	1	28.03333333	50.61574074	0.000000003250293
Within Groups	28.8	52	0.5538461538		
Total	56.83333333	53			

It can be generalized that the chatbot intervention based on the proposed model had a positive effect on the results of the test group compared to the control group that did not use the chatbot, in terms of eating habits, finding alternatives, and progressing in the phases of the new model. However, patients still have difficulties committing to the treatment plan, which may be caused by the difficulty of the treatment plan itself, which includes stopping eating many of the common meals consumed in daily life and social occasions. Also, the reason for the decrease in chatbot usage rate from month to month may be due to the fact that the communication with the chatbot is by using writing, which allows reading the messages again and again that makes the user benefit from the chatbot message history without making a new chat with the chatbot.

VIII. Conclusion

This study built a new health behavior change model based on the common existing models. Then, it used that model to design and test a chatbot for celiac patients in order to find out the answers to the research questions: Does the proposed behavior-changing model-based design of WhatsApp chatbot help celiac patients improve their eating habits?

In order to do that, the data were collected using three methods: studying the most common health behavior change models (HBM, TPB, DOI, TTM, and SNT), interviews with patients, their parents, dietitians, and gastroenterologists, and questionnaires for celiac patients and their relatives. Using these data, the new model was built (called CDEM) and adapted for use in technology. As a proof of concept, the CDEM was applied to WhatsApp chatbot for patients with celiac disease.

Two groups participated in the experiment: the control group and the test group. The control group was left without any intervention, whereas the test group used the chatbot. The results showed improvement in the eating habits of patients with celiac disease after using the chatbot in the test group compared to the control group. Therefore, it can be said that the proposed behavior-changing model-based design of WhatsApp chatbot can help celiac patients improve their eating habits.

This study faced many challenges because the timing of this study coincided during the Covid-19 pandemic, which led to limited communication with the target group, as well as limiting the pre-test and post-test to electronic tests. The study was conducted in the Kingdom of Saudi Arabia and on a Saudi sample from Jeddah only. The same study may provide different results with a different demographic sample or a diverse demographic group.

Our future direction is to conduct more experiments on the CDEM and use it in different social networks, since the experiment was applied in WhatsApp only. In the future work, the authors will study another social network apps' ability to change behavior, as we believe that some apps will give better results than others in achieving the goal. Also, may be carried out on other categories of society to improve the quality of life for sustainability factors.

References

[1] S. Kemp, "Digital 2021 global overview ", We are Social, 2021.

[2] K. Okeleke and S. Suardi, "The mobile economy 2021", GSMA Intelligence, Jun 2021.

[3] A. Smith and D. Page, "US smartphone use in 2015," 2015.

[4] M. Peleg et al., "Ideating mobile health behavioral support for compliance to therapy for patients with chronic disease: a sase study of Atrial Fibrillation management," Journal of Medical Systems, vol. 42, no. 11, p. 234, 2018/10/13 2018, doi: 10.1007/s10916-018-1077-4.

[5] D. Y. P. Chao, T. M. Y. Lin, and W.-Y. Ma, "enhanced self-efficacy and behavioral changes among patients with diabetes: cloud-based mobile health platform and mobile app service," JMIR Diabetes, vol. 4, no. 2, p. e11017, 2019/05/10 2019, doi: 10.2196/11017.

[6] K. Regmi, N. Kassim, N. Ahmad, and N. Tuah, "Effectiveness of mobile apps for smoking cessation: α review," Tob. Prev. Cessation, vol. 3, 04/12 2017, doi: 10.18332/tpc/70088.

[7] D. M. Kazemi, B. Borsari, M. J. Levine, K. A. Lamberson, and B. Dooley, "REMIT: Development of a mHealth theory-based intervention to decrease heavy episodic drinking among college students," Addiction Research & Theory, vol. 26, no. 5, pp. 377-385, 2018/09/03 2018, doi: 10.1080/16066359.2017.1420783.

[8] D. J. Dute, W. J. E. Bemelmans, and J. Breda, "using mobile apps to promote a healthy lifestyle among adolescents and students: a review of the theoretical basis and lessons learned," JMIR mHealth uHealth, vol. 4, no. 2, p. e39, 2016/05/05 2016, doi: 10.2196/mhealth.3559.

[9] S. R. Paige, J. M. Alber, M. L. Stellefson, and J. L. Krieger, "Missing the mark for patient engagement: mHealth literacy strategies and behavior change processes in smoking cessation apps," Patient Education and Counseling, vol. 101, no. 5, pp. 951-955, 2018/05/01/ 2018, doi: https://doi.org/10.1016/j.pec.2017.11.006.

[10] A. Annie, "The state of mobile.," 2019.

[11] J. Åberg, "Chatbots as a mean to motivate behavior change : how to inspire pro-environmental attitude with chatbot interfaces," Independent thesis Advanced level (degree of Master (Two Years)) Student thesis, 2017. [Online]. Available: http://urn.kb.se/resolve?urn=urn:nbn:se:umu:diva-135835.

[12] WhatsApp. "Coronavirus - WHO health alert launch " WhatsApp.com. https://www.whatsapp.com/coronavirus/who/ (accessed 11 April, 2020).

[13] Telegram. "Coronavirus news and verified channels." Telegram.com. https://telegram.org/blog/coronavirus (accessed 19 April], 2020).

[14] E. Lionetti and C. Catassi, "New clues in celiac disease epidemiology, pathogenesis, clinical manifestations, and treatment," International Reviews of Immunology, vol. 30, no. 4, pp. 219-231, 2011/07/29 2011, doi: 10.3109/08830185.2011.602443.

[15] P. Singh et al., "Global prevalence of Celiac disease: systematic review and meta-analysis," Clinical Gastroenterology and Hepatology, vol. 16, no. 6, pp. 823-836.e2, 2018/06/01/ 2018, doi: https://doi.org/10.1016/j.cgh.2017.06.037.

[16] Y. M. Khayyat, "Serologic markers of gluten sensitivity in a healthy population from the western region of Saudi Arabia," (in eng), Saudi J Gastroenterol, vol. 18, no. 1, pp. 23-25, Jan-Feb 2012, doi: 10.4103/1319-3767.91733.

[17] A. Al-Hussaini et al., "Mass screening for celiac disease among school-aged children: toward exploring celiac iceberg in Saudi Arabia," Journal of Pediatric Gastroenterology and Nutrition, vol. 65, no. 6, 2017. [Online]. Available: https://journals.lww.com/jpgn/Fulltext/2017/12000/Mass_Screening_for_Celiac_Disease_Among.12.aspx.

[18] M. A. Safi, "Prevalence of Celiac disease in Saudi Arabia: meta-analysis," Global Vaccines and Immunology, vol. 3, 01/01 2018, doi: 10.15761/GVI.1000134.

[19] J. L. Nagel, "Survey of registered dietitians' proficiency of Celiac disease and use of Twitter, Facebook, Smart Phone app, and internet for Celiac Disease management," ed: North Dakota State University, 2014.

[20] A. J. Dowd, C. Jackson, K. T. Y. Tang, D. Nielsen, D. H. Clarkin, and S. N. Culos-Reed, "MyHealthyGut: development of a theory-based self-regulatory app to effectively manage celiac disease," (in eng), Mhealth, vol. 4, p. 19, 2018, doi: 10.21037/mhealth.2018.05.05.

[21] K. Park, M. Harris, N. Khavari, and C. Khosla, "Rationale for using social media to collect patient-reported outcomes in patients with Celiac

Disease," (in eng), J Gastrointest Dig Syst, vol. 4, no. 1, p. 166, 2014, doi: 10.4172/2161-069X.1000166.

[22] A. M. H. Alshebani and Z. S. Abdalhamza, "The role of celiac disease antibodies in the follow up of patient on gluten free diet " Revista Latinoamericana de Hipertension, vol. 13 6, pp. 561-566.SEP, 2018.

[23] G. Midhagen et al., "Antibody levels in adult patients with coeliac disease during gluten-free diet: a rapid initial decrease of clinical importance," Journal of Internal Medicine, vol. 256, no. 6, pp. 519-524, 2004, doi: 10.1111/j.1365-2796.2004.01406.x.

[24] I. M. Rosenstock, V. J. Strecher, and M. H. Becker, "Social learning theory and the health belief model," Health Education Quarterly, vol. 15, no. 2, pp. 175-183, 1988/06/01 1988, doi: 10.1177/109019818801500203.

[25] M. Conner, "Theory of planned behavior," Handbook of Sport Psychology, p. 3, 2020.

[26] G. Silverberg, G. Dosi, and L. Orsenigo, "Innovation, diversity and diffusion: a self-organisation model," The Economic Journal, vol. 98, no. 393, pp. 1032-1054, 1988, doi: 10.2307/2233718.

[27] The transtheoretical model: Applications to exercise. (1994). Human Kinetics Publishers, Champaign, IL, England.

[28] H. W. Perkins and A. D. Berkowitz, "perceiving the community norms of alcohol use among students: some research implications for campus alcohol education programming*," International Journal of the Addictions, vol. 21, no. 9-10, pp. 961-976, 1986/01/01 1986, doi: 10.3109/10826088609077249.

[29] K. Sainsbury and B. Mullan, "Measuring beliefs about gluten free diet adherence in adult coeliac disease using the theory of planned Behavior," Appetite, vol. 56, no. 2, pp. 476-483, 2011/04/01/ 2011, doi: https://doi.org/10.1016/j.appet.2011.01.026.

[30] C. Shao, J. Wang, J. Liu, F. Tian, and H. Li, "Effect of a Health Belief Model-based education program on patients' belief, physical activity, and serum uric acid: a randomized controlled trial," (in eng), Patient Prefer Adherence, vol. 12, pp. 1239-1245, 2018, doi: 10.2147/PPA.S166523.

[31] L. H. McArthur, A. Riggs, F. Uribe, and T. J. Spaulding, "Health Belief Model offers opportunities for designing weight management interventions for college students," Journal of Nutrition Education and Behavior, vol. 50, no. 5, pp. 485-493, 2018/05/01/ 2018, doi: https://doi.org/10.1016/j.jneb.2017.09.010.

[32] B. Mohebbi, A. Tol, R. Sadeghi, S. F. Mohtarami, and A. Shamshiri, "Self-management intervention program based on the Health Belief Model (HBM) among women with Gestational Diabetes Mellitus: a quazi-experimental study," (in eng), Arch Iran Med, vol. 22, no. 4, pp. 168-173, Apr 1 2019.

[33] A. K. Jeihooni, S. F. Dindarloo, and P. A. Harsini, "Effectiveness of Health Belief Model on oral cancer prevention in smoker men," Journal of Cancer Education, vol. 34, no. 5, pp. 920-927, 2019/10/01 2019, doi: 10.1007/s13187-018-1396-7.

[34] R. A. Record, N. G. Harrington, D. W. Helme, and M. W. Savage, "Using the Theory of Planned behavior to guide focus group development of messages aimed at increasing compliance with a tobacco-free policy," (in eng), Am J Health Promot, vol. 32, no. 1, pp. 143-152, Jan 2018, doi: 10.1177/0890117116687467.

[35] İ. Dilekler, C. Doğulu, and Ö. Bozo, "A test of theory of planned behavior in type II diabetes adherence: The leading role of perceived behavioral control," Current Psychology, 2019/05/27 2019, doi: 10.1007/s12144-019-00309-7.

[36] F. Abamecha, A. Tena, and G. Kiros, "Psychographic predictors of intention to use cervical cancer screening services among women attending maternal and child health services in Southern Ethiopia: the theory of planned behavior (TPB) perspective," BMC Public Health, vol. 19, no. 1, p. 434, 2019/04/25 2019, doi: 10.1186/s12889-019-6745-x.

[37] H. Trisnowati, D. Kusuma, A. Ahsan, D. E. Kurniasih, and R. S. Padmawati, "Smoke-free home initiative in Bantul, Indonesia: Development and preliminary evaluation," (in eng), Tob Prev Cessat, vol. 5, pp. 40-40, 2019, doi: 10.18332/tpc/113357.

[38] L. Paina et al., "Applying the model of diffusion of innovations to understand facilitators for the implementation of maternal and neonatal health programmes in rural Uganda," Globalization and health, vol. 15, no. 1, p. 38, 2019, doi: 10.1186/s12992-019-0483-9.

[39] A. Koyun and K. Eroglu, "The effect of transtheoretical model-based individual counseling, training, and a 6-month follow-up on smoking cessation in adult women: A randomized controlled trial," TURKISH JOURNAL OF MEDICAL SCIENCES, vol. 46, pp. 105-111, 01/05 2016, doi: 10.3906/sag-1407-100.

[40] E. Shakiba et al., "Efficacy of transtheoretical model on preventive nutritional behaviors of cardiovascular diseases: a randomized controlled trial," J-Mazand-Univ-Med-Sci, vol. 28, no. 163, pp. 24-37, 2018. [Online]. Available: http://jmums.mazums.ac.ir/article-1-10713-en.html.

[41] R. d. M. Boff, M. A. Dornelles, A. M. P. Feoli, A. d. S. Gustavo, and M. d. S. Oliveira, "Transtheoretical model for change in obese adolescents: MERC randomized clinical trial," Journal of Health Psychology, p. 1359105318793189, 2018, doi: 10.1177/1359105318793189.

[42] M. Cooper, M. Creamer, C. Ly, B. Crook, M. Harrell, and C. Perry, "Social norms, perceptions and dual/poly tobacco use among texas youth," American journal of health behavior, vol. 40, pp. 761-770, 01/11 2016, doi: 10.5993/AJHB.40.6.8.

[43] P. Pansu, V. Fointiat, L. Lima, C. Blatier, P. Flore, and N. Vuillerme, "Changing behaviors: using norms to promote physical activity for type 2 diabetes patients," European Review of Applied Psychology, vol. 69, no. 2, pp. 59-64, 2019/03/01/ 2019, doi: https://doi.org/10.1016/j.erap.2019.03.001.

[44] S. Higgs, J. Liu, E. I. M. Collins, and J. M. Thomas, "Using social norms to encourage healthier eating," Nutrition Bulletin, vol. 44, no. 1, pp. 43-52, 2019/03/01 2019, doi: 10.1111/nbu.12371.

[45] J. Finkelstein and E. M. Cha, "Using a mobile app to promote smoking cessation in hospitalized patients," JMIR mHealth uHealth, vol. 4, no. 2, p. e59, 2016/05/06 2016, doi: 10.2196/mhealth.5149.

[46] C. Khalil, "Understanding the adoption and diffusion of a telemonitoring solution in gestational diabetes mellitus: qualitative study," JMIR Diabetes, vol. 4, no. 4, p. e13661, 2019/11/28 2019, doi: 10.2196/13661.

[47] N. Alsaleh and R. Alnanih, "Gamification-based behavioral change in children with diabetes mellitus," Procedia Computer Science, vol. 170, pp. 442-449, 2020/01/01/ 2020, doi: https://doi.org/10.1016/j.procs.2020.03.087.

[48] H.-y. Shum, X.-d. He, and D. Li, "From Eliza to XiaoIce: challenges and opportunities with social chatbots," Frontiers of Information Technology & Electronic Engineering, vol. 19, no. 1, pp. 10-26, 2018/01/01 2018, doi: 10.1631/FITEE.1700826.

[49] F. Dubosson, R. Schaer, R. Savioz, and M. Schumacher, "Going beyond the relapse peak on social network smoking cessation programmes: ChatBot opportunities," Swiss Medical Informatics, 09/20 2017, doi: 10.4414/smi.33.00397.

[50] J. Beaudry, A. Consigli, C. Clark, and K. J. Robinson, "Getting ready for adult healthcare: designing a chatbot to coach adolescents with special health needs through the transitions of care," Journal of Pediatric Nursing, vol. 49, pp. 85-91, 2019/11/01/ 2019, doi: https://doi.org/10.1016/j.pedn.2019.09.004.

[51] P. Lally, C. H. M. van Jaarsveld, H. W. W. Potts, and J. Wardle, "How are habits formed: modelling habit formation in the real world," European Journal of Social Psychology, vol. 40, no. 6, pp. 998-1009, 2010, doi: 10.1002/ejsp.674.

[52] M. Maltz, New Psycho-Cybernetics. Penguin, 2002.

[53] M. Gladwell, Outliers: the story of success. Audio-Tech Business Book Summaries, 2017.

[54] S. Terrell, and D. Miranda, The Transtheoretical Model (TTM, the. Health Care Financing Review 2001, 23(1), p.87.