# How to Write Health Dialog for a Talking Computer

**Jeffrey P. Migneault, Ramesh Farzanfar, Julie Wright, Robert H. Friedman**

**Medical Information Systems Unit, Department of Medicine**
**Boston University and Boston Medical Center**

Corresponding Author: Jeffrey P. Migneault Medical Information Systems Unit. 560 Harrison Avenue, Suite 405 Boston, MA 02118 jpm@bu.edu. Fax: (617) 638-5580

Abstract

Automated dialogue systems delivered over the telephone offer a promising approach to delivering health-related interventions to populations of individuals at low-cost. Over the past two decades, an automated telephone system called Telephone Linked Care or TLC has been successfully designed and evaluated by the authors and their colleagues. This work has resulted in over twenty systems for various health-related conditions and lifestyle behaviors. This paper describes our approach to developing and writing dialogue for these automated telephone systems, including determining the program objectives, defining the target population, and selecting a theory of behavior change to guide the intervention. Both macro and micro issues are considered in constructing dialogue systems that are engaging for the target population, easy to use, and effective at promoting positive health behaviors and outcomes.

**Keywords: dialogue systems, behavioral change systems, telephone linked care**

## 1. Introduction

Automated dialogue systems are increasingly being used in health care to provide information, advice, counseling, disease monitoring, clinical problem identification, as well as enhancing patient-provider communication. They are also being used with the general population (consumers) to improve health related lifestyle behaviors. We have 20 years of experience developing and evaluating completely automated telephone-based conversational systems that interact with patients and consumers to improve health outcomes and the delivery of health services. In this paper, we will present our experiences designing these systems, which we call Telephone-Linked Care or TLC systems. We will describe the process of developing conversational structure and content as well as the process of implementation.

The systems we have developed and evaluated have focused on (1) positively influencing a person's health behavior by modifying behavioral risk factors for disease (e.g., smoking, diet quality, etc.) and promoting disease-related self-care behaviors (e.g., taking prescribed medications regularly, attending scheduled clinical office visits, etc.), and (2) monitoring patient's health conditions outside of clinical settings and identifying potential acute, sub-acute, and chronic medical care issues that are then communicated to the person's responsible health professional(s). For this paper, our focus will be on the dialogues necessary for the first of these objectives, namely improving health behaviors of patients and consumers, although most of the considerations also apply to the second objective of disease monitoring, an integral component of disease management programs. For the improvement of lifestyle behaviors, we have developed TLC systems for several aspects of diet [1-3], physical activity [4-6], and cigarette smoking [7]. For disease related self-care behaviors, there are TLC programs for promoting medication-regimen adherence [3, 8-10], scheduled office visit attendance [10], appropriate disease screening behavior [11] and use of home self-measurement devices [8]. TLC chronic disease management programs exist for hypertension [8], angina pectoris, chronic obstructive lung disease [12, 13], asthma [14], diabetes mellitus, depression [10], and patients with multiple chronic diseases.

Over time our program objectives have evolved from being single behavioral objectives (e.g., improve medication adherence) to multi-behavioral interventions (medication adherence, diet, and exercise). This evolution reflects the increasing functionality of TLC programs and a recognition that improving health or self-care of chronic conditions (e.g. hypertension) is a multi-behavioral enterprise. Evaluations of TLC systems have demonstrated efficacy in improving target health behavior in multiple studies. Evaluations of TLC systems have compared TLC to "usual care" and to mock-interventions that used TLC technology to control for the attention-placebo effect of the contact itself and the novelty of the TLC intervention. Using the standard metric of effect sizes (behavior change expressed in terms of standard deviations on the outcome variables), TLC systems have demonstrated medium to large effect sizes ranging from .29 to .69 [15]. More recently, a TLC physical activity promotion program was compared to both an assessment-only control and a human telephone counselor program [16]. The results demonstrated that both the human and TLC groups were effective compared to the control and were not significantly different from each other.

## 2. Automated Telephone Systems or Telephony

The use of the telephone as the communication channel of the health behavior change programs has a number of advantages. The telephone is almost universally available in the U.S., and in both developing and developed countries. In fact, with the spread of mobile telephone technology globally, many developing countries have built national mobile telephone systems while their landline telephone systems remain rudimentary. Because telephones are ubiquitous and have been a mainstay of communication between people, typically two people at a time, using a phone for conversation with another person is perceived as natural and easy. Our research has shown that users actively project human qualities onto the automated voice and this lends credibility to the system [17, 18]. Although it

has not been empirically tested, we believe that this projection is more likely to occur in telephone encounters than in programs that use other technological channels of communications. Interestingly, this effect might not be entirely positive as in some situations, individuals may be more honest with less personal (computers vs. live interviewers) assessment methods[19, 20].

Unlike human-to-human interactions, the TLC system's conversations with patients and consumers are entirely automated. During TLC telephone conversations, the system speaks to individuals over the telephone using digitally recorded human voices. Users communicate with TLC by pressing buttons on their telephone keypad, or in newer systems, by speaking into the telephone receiver.

The interaction is designed to resemble a typical conversation between a health counselor and a patient/client. The use of speech recognition technology helps make the interaction "feel" like a natural conversation. At the core of these systems is the dialogue, the exact words that the system 'speaks' to the user during TLC conversations. We refer to the written version of this dialogue as the 'script.' The script contains the conversational segments to be digital recorded by voice consultants (actors) into separate sound files as well as the rules for assembling these segments into the resulting, often unique, conversation. This includes any questions that are going to be asked during the conversation and the possible answers that the system will accept. How skillfully the dialogue is designed and written largely determines the nature and effectiveness of telephony systems.

Our current TLC programs are comprised of a computer system that combines an interactive voice response (IVR) subsystem for generating speech using prerecorded audio message segments, a speech recognition subsystem for recognizing what the user is saying, a database management subsystem for storing and managing system and user data, and a conversation control subsystem that controls the content and flow of individual TLC conversations with users. TLC systems are finite state machines in which programmed decision rules are used to select, combine and play pre-recorded sound files of conversational segments. These rules determine which sound files to play based on logic that is applied at each step of the conversation and data the system has about the user and their progression through the program. The data on user characteristics might be known before a contact is initiated or collected from the user during the call.

We design TLC systems so that they are tailored to the user's 'state' using current behavioral theories to structure the design of the systems and the strategies used to promote change. Since behavioral theories do not usually delineate the specific tactics that should be used by the systems to affect the behavioral strategies, we incorporate the heuristics of experienced clinicians in the tactical design and in crafting the specific words uttered by TLC. Given the complexity of these design considerations, the variety of component behaviors a program might be addressing, and the number of user states given the factors we are tailoring on, TLC systems are quite large and complex. The printed representations of TLC 'scripts' are usually in the range of 150-600 pages. We know of no other evaluated health dialogue systems that approach TLC programs in their size, breadth, depth, complexity of content or in the degree of incorporation of behavioral change theories and heuristics of health professionals. This paper will describe the most salient attributes and our process for developing these systems.

## 3. Initial Considerations

Before beginning to write dialogue for health behavior change systems, a number of parameters should be carefully examined and clearly delineated, including the objectives of the program, the characteristics of the population that will be using the system, and the intervention approach that will be employed. Having these parameters well defined will make the development process more efficient and lessen the chance that changes will be needed after the program has been initiated, which is quite costly for automated systems.

3.1 Program Objectives

Although any intervention program will at the outset have an overall objective, it is important for an automated health behavior intervention program to have objectives that are carefully considered and explicitly delineated. The program's objectives need to be kept in mind throughout the design and writing phases of intervention development. Adjustments to the objectives often occur during the developmental process and should be explicitly documented. This process is imperative when multiple developers are working on the system. The objectives should be defined not just in terms of the overall goal (e.g., promoting physical activity) but include: (1) what aspects of a behavior or behaviors will be changed, (2) what are the ultimate goals and what are the intermediate goals, (3) in what order will the goals be accomplished, and (4) how will the achievement of the goals be assessed.

The first of these considerations is important because most lifestyle behaviors are quite complex and any one program is likely to focus on only some aspects of the behavior. For example, in one of our current exercise programs we are specifically targeting levels of moderate intensity aerobic activity such as brisk walking. We are not trying to improve strength, flexibility, or promote vigorous intensity exercise although incidental improvements in these will certainly not be discouraged. In this same program, specifications of ultimate and intermediate goals (#2 from above) would include the ultimate goal of getting users to reach and maintain recommended levels of moderate or greater intensity exercise of 30 minutes per day most days of the week (or 150 minutes per week) [21]. Intermediate goals are often different for different users depending on their characteristics. For example, intermediate goals may include increasing the level of motivation for users with low motivation, developing specific plans for those users who are motivated but do have not a viable exercise plan, increasing total minutes of exercise for those who are getting less than 150 minutes per week in small but steady increments, and preventing relapse for those who achieve this level of exercise. The third consideration, the order of goals, is especially important in programs that target more than one health behavior, but can also be important when intermediate goals do not have an inherent order. The last aspect of the objectives to be defined (#4) is to specify how they are going to be measured. We assess the targeted behavior throughout the program for two reasons: First, to provide data to individualize intervention messages as they are being delivered, and, second, for program evaluation. Assessment procedures for these two purposes can be independent or overlapping. There are a number of study design issues to be considered, but this topic is beyond the scope of this paper.

3.2 Population Characteristics

The second important parameter is understanding the characteristics of the population that will use the program. This knowledge is necessary so that the dialogue will be engaging to all or nearly all users. Furthermore, as an integral aspect of these systems is to individualize or tailor the intervention to each user, knowledge of the population can assist in identifying the critical individual and group differences to take into account in dialogue construction. A non-exhaustive set of important characteristics to consider in dialogue writing include education level, gender, age, culture or ethnicity, income, and geographic characteristics of the home neighborhood. Since the dialogue is designed for the whole population, understanding the range of values on a characteristic is more important than knowing the mean value for the population. For example, knowing the prevalence of low educational levels in a target population for an automated health behavior program is more helpful than understanding what the average level is. Interventions adjusted for low education level (or low health literacy) are generally effective even for those with higher levels of education. That said, decisions also might be made not to adapt the dialogue to the extremes of a dimension. If 98% of the population has at least a sixth grade education, it might not make sense to adapt the intervention to the 2% of the

population that has lower education levels. In this case, a decision would need to be made whether or not to exclude individuals with this level of education.

A new area of research for our group is investigating interventions tailored to specific racial or ethnic groups [3]. The ethnic identity of the user provides the primary context in which the lifestyle behaviors occur. For instance, one's culture has a large effect on dietary preferences and the role of food in self-concept and social interactions. In our work on culture, we use Resnicow and colleagues' framework [22], which emphasizes understanding both 'surface' aspect of culture (e.g., food preferences) as well as 'deep' aspects (e.g., role of food in maintaining interpersonal relationship).

In addition to accessing data on distributions across variables, we have found that two other sources provide invaluable information on populations we serve. The first is direct experience, usually in clinical settings, with members of the target population. If this experience does not exist with the design team, it is advisable to find outside consultants to review program design and the dialogue writing approach. The second source of information is from focus groups of individuals from the target population. In an ongoing project to improve self-care for hypertensive African-Americans, we ran a series of focus groups on self-care behaviors, relationships with care givers and the experience of hypertension. We listened carefully not only to what was said but to how it was said. Idiosyncratic terminology used by a sub-group can be used to adapt the intervention scripts. Once an in depth understanding of the population is achieved, the writing of the dialogue will be better informed producing more effective and engaging interactions.

### 3.3 Intervention approach: Using theory to guide the development of a dialogue system

The third set of decisions to be made concerns the intervention approach that will be used to change the behavior. The chosen approach should be explicitly stated and used as a guide to dialogue writing. We believe the most important of these intervention approach decisions is selecting a behavioral change theory to guide the construction of the dialogue system.

Some of the most effective health behavior change programs use theory as a framework for designing and implementing the intervention. Theory is defined as a set of interrelated concepts, definitions, and propositions that present a systematic view of events or situations by specifying relations among variables in order to explain and predict these events and situations [23, 24]. In health behavior research, theory is essentially a framework that determines the approach to and the components of the intervention. There are numerous theories available to researchers building health behavior interventions, e.g., Social Cognitive Theory [25, 26], Health Belief Model [27], Theory of Reasoned Action [28], Theory of Regulation and Self-control [29], and Theory of Subjective Culture and Interpersonal Relations [27-31], Transtheoretical Model of Behavior Change [32-34], Protection Motivation [35], and Precaution Adoption Process [36]. Our laboratory has structured the majority of our health behavior programs using established behavioral theories including the Transtheoretical Model and Social Cognitive Theory (SCT). The purpose of this paper is not to propose the use of any one specific theory to guide the dialogue system; our purpose is to recommend that dialogue systems for health behavior change be based on a theory. We believe this is important for two reasons. First and most immediately, it is likely to lead to more useful programs because good theory leads to more comprehensive and consistent intervention designs and provides effective strategies for behavior change. The second reason is that evaluating theory-based programs provides important data to the field on how to improve theories of health behavior change which ultimately lead to more effective interventions.

In choosing a theoretical basis for a specific dialogue system, the system developer should evaluate candidate theories on a few main criteria that we have found to be important. The first is whether the theory has any empirical evidence related to improving the target health outcomes. The second is to determine if the theory has a framework that it is easily translatable into a dialogue

conversation. Since TLC conversations are aimed at the individual user, it is important that the theory is able to explain an individual's behavior as opposed to explaining population or group behavior. Thirdly, the theory should provide clear guidance on the timing and content of the intervention, i.e., what to say and when to say it, so computer algorithms to control the conversation can be specified.

One theory that we have used many times in designing our TLC automated dialogue systems is the Transtheoretical Model of Behavior Change (TTM). The TTM integrates a set of constructs related to how individuals change a health-related behavior; a detailed description of the model can be found in an article by Prochaska and Velicer (1997) [37]. From our perspective, the TTM meets all three criteria delineated above. First, there is extensive empirical evidence on the effectiveness of the TTM for health behavior change [32, 33]. The precepts of the TTM model have been used to structure a variety of health behavior change intervention programs, and these programs have been shown in well designed evaluation studies to change health behavior in the ways predicted by the model. Second, as the theory was developed to specifically explain the process of behavior change, its conceptual structure facilitates its translation from a theory to an intervention. The theory consists of four interrelated constructs, each of which is easily translated into dialogue. The third quality that makes the TTM an appropriate basis for an automated intervention is that it provides guidance on what content is to be delivered and in what situations; thus, computer algorithms can be easily rendered.

3.4 An Example of How to Use Theory to Guide the Development of a Dialogue System.

The following section uses information on one of our physical activity programs (TLC-PA), developed using the TTM as a theoretical framework as an example of what a dialogue system might include. Based on current health recommendations, a typical goal of our physical activity programs is to assist users to achieve and maintain 150 or more minutes per week of moderate-intensity physical activity [21]. Four TTM constructs were used to design TLC-PA, namely: stage of change, processes of behavior change, decisional balance, and self-efficacy. The first of these represents an individual's progression through a change process and provides an organizing structure for the integration of the other three constructs, therefore, we designed TLC-PA to provide separate dialogues for each of the five stages of change: Precontemplation, Contemplation, Preparation, Action and Maintenance [38]. The user's stage is assessed at the beginning of each call through a series of three to five branched questions. If they have changed stage since the previous call, they are given appropriate feedback on this movement in stage. The caller will then hear content based on the other three theoretical constructs but tailored both by their current stage of change and what content they have progressed through previously. For example, a person in the Contemplation Stage would engage in a conversation to clarify their decision to exercise or not (the decisional balance construct) and one part of this would might be about the many benefits of doing regular moderate physical activity and including: "*O.K., I will read a list of several benefits to you. You can pick the topic most important to you and I will then tell you how exercise can help*." The processes of change are a set of overt (actions) or covert (thinking/feeling) activities that foster positive behavior change and are employed in all stage dialogues. Which process is most beneficial at any one time is largely determined by which stage the user is in. For example, the process of consciousness raising might be used in the Preparation stage by encouraging a user to learn more about the types of physical activity they might consider. The conversation for this user might then move on to building the user's confidence that they can successfully engage in regular exercise (the construct of self-efficacy). The introduction to this section includes: "A*s you move to becoming and staying regularly active, there may be times when you have doubts whether you can make these changes. If you would like to hear how to feel more confident about being active, say 'confident', if you'd like to skip ahead, say 'skip' now."* Finally, this conversation would end with TLC-PA negotiating a stage-specific goal for the next week. For instance, for those in Preparation who have begun to increase their level of physical activity, the system has a set of

algorithms that considers the overall physical activity goal of the program, the person's previously stated intermediate goals, the person's previous levels of physical activity and heuristics that consider appropriate changes in goals over time. For example, after asking the caller about their goal, TLC might say, "*Exercise experts and physicians agree that people can and should slowly increase their amount of exercise activity; even if it is only by 5 minutes each day. Previously, you decided to exercise on [previous planned days]* days for [previous planned minutes]* minutes each day. I think it's reasonable for you to try to do 5 minutes more exercise, on each day that you exercise this coming week. Would you like to re-consider your activity plans for the coming week? Say 'Yes' or 'No' now."* Alternatively, if the user tries to set an intermediate goal that is much higher than their current exercise level, the TLC system will encourage them to adjust their next goal downward to a level where success is more likely. This strategy process is in keeping with the theory and empirical findings on self-efficacy, namely, that it is increased by success on intermediate goals. In summary, the TTM was used throughout the design of this intervention, it use made success more likely, and studies using this system are a good test of the usefulness of this theory in the design of effective interventions.

(footnote to go with *: [previous planned days] and [previous planned minutes] represent variables that contain the number of days and minutes the users chose as their plan during the previous TLC call one week earlier.)

**4. System Specifications**

One of the next steps in the process of designing an automated TLC program is determining the systems specifications. Specifications include the number of contacts, the duration of the program, the schedule of contacts, and the duration of individual TLC conversations. These attributes are specific to each application and are determined by the objectives, population characteristics, and the demands of the intervention approach but are also tempered by practical considerations. The total contact time designed into a program is determined by these parameters with each dimension being independently important. TLC programs have varied greatly on all three of these dimensions.

4.1 Number of User-System Contacts

We have developed TLC systems for a range of contacts, from as few as a single contact to systems that can handle an indefinite number of contacts. An example of a system designed for one contact is our TLC program to promote screening mammography in women. This system contacts the woman a month before her annual mammogram is due. The one-call design fits the behavioral objective(to have the woman take a single action) and is sufficient to present the content hypothesized to bring about this effect. The intervention both assists the user directly in the scheduling of the screening and addresses whatever barriers to getting a mammogram that the woman might have. The system allows the user to choose from a set of 22 barriers to hear advice that is applicable to her. The system also allows the women to call back if they were interested in hearing more barriers or wanted information repeated. In a recent study we found that few women utilized this call back option which was indirect support for the 1-contact design [11]. Also the mean call duration was 11½ minutes, and the range of durations was from 3½ to 27 minutes indicating that women chose to hear very different amounts of content.

Given that many health behaviors (e.g., eating, being physically active/inactive, drinking alcohol, smoking cigarettes, etc.) are inherently complex and often resistant to change, they require much more intensive programs than the mammography program just described. The majority of our TLC systems have used multiple contacts over time. In fact, we have designed systems with the potential to provide on-going contacts for an indefinite amount of time in order to foster life-time maintenance of health behaviors.

4.2 Duration of the Program

We have found that the duration of a behavioral program should be long enough for users to receive the full content of the program at a reasonable rate for them, have time to practice and integrate new skills or perspectives into their lifestyle both between and across contacts, have the opportunity to repeat difficult sections of the program, and achieve success and maintain it for at least a short period of time. The majority of our programs to date have been designed to be delivered over a six month period which has resulted in positive health outcomes [16]. One of our TLC physical activity programs was delivered over the course of a year. It had positive effects after six months but there were no further increases between 6 and 12 months [16], suggesting that six months of TLC was sufficient at least for initiation of behavior change.

Most of our work to date has been in the initiation of behavior change, whereas maintenance of behavior change might call for different program designs, such as those with durations longer than a year or those which are provided on an indefinite basis As mentioned earlier, one of the outstanding advantages of automated systems is that once developed, the incremental cost of use is low. Unfortunately, few data exist to inform design decisions in this area given that few studies have systematically investigated the use of automated systems to promote the long term maintenance of health behavior change. We are presently completing two TLC systems that expressly target the long-term maintenance of dietary improvement, a very difficult problem. These two systems are based on completely different theories of behavior change, and they will be compared both to each other and to a control group. Although dialogue systems allow for program designs of indefinite length, more research is needed to determine the optimal duration for the various objectives of these systems.

4.3 Schedule of Contacts

Not only does the duration of a program vary, but the schedule of TLC contacts within this time period varies as well. Although the most common frequency of contacts is one call per week, other options are possible. One clear advantage of regular (daily or weekly) calls is that it is easier for a user to get into the habit of having a call at that time, improving call adherence. We are currently involved in the design of a system that will use very short daily calls to just gather information on two behaviors (physical activity and glucose testing), and then this information will be combined and used in the more usual weekly behavior change conversations. Other TLC programs start off with weekly calls, and then after the main content has been completed, reduce frequency to bi-weekly and then monthly calls that become more of a check-in and review than presentation of new content. This variation in call frequency has lead to six month programs with a low of nine contacts (weekly for one month, followed by monthly for five months) to those with 26 weekly calls . Again, these decisions are made based on the amount of content the program needs to cover and our understanding of what might be optimal given the program objectives and population of users. However, we have not directly compared outcomes of programs with different durations or contact frequency.

4.4 Duration of TLC Conversations

The duration of each TLC conversation varies across the programs, across different calls within a program, and among the users of a particular program. The average duration of TLC calls was as short as four minutes for the original TLC-Hypertension system [8] and as long as 25 minutes for an intervention to increase physical activity. In general, we have found that calls of 10–15 minutes tend to be optimal for most users. Given that it usually takes a couple of minutes to initiate the call and end the call with a short review and confirmation of the next contact time, there is between 8-13 minutes for the main content of the intervention per call. In some systems, the decision whether or not to go on to a new section is made during the call. If the elapsed time is past a limit, say 10 minutes, the system

might decide to not start a new section or, alternatively, ask the user if they want to go on or start the section in the next call. This takes some extra programming but can be a nice way to give the user more control over the conversation.

We have limited the duration of individual TLC calls for a number of reasons. First, we wished to minimize the users' perceived burden as it is likely to be associated with attentiveness during the calls, rates of call completion, and adherence to the prescribed use schedule. Second, we wished to not overload the user, but rather maximize the likelihood that the user comprehends and retains the information and counseling delivered. Often when adapting a non-automated intervention to a dialogue system, the content of what might be delivered in an ideal assessment and intervention contact will have to be divided into the smallest viable units and delivered across multiple TLC conversations. For example, the total amount of time to deliver one complete dietary assessment and intervention in our TLC "healthy eating" dietary program [1] would be about one hour. We divided the material into six conversations of about ten minutes each, covering all of the content. These six conversations were called a cycle. When one cycle was completed, there was a review call, and then a new cycle of calls would begin.

4.5 Other Important System Specifications

There are other design decisions that although more limited in scope, can have large effects on intervention effectiveness. These specifications include how the user-TLC conversation is initiated, and the automated options related to user-system interactions and response options.

Initiation of the TLC Conversation. Automated telephony technology permits both inbound and outbound calls, or in other words, either the user can call the system (inbound), or the system can call the user (outbound). From a programming perspective, the inbound contact systems are much simpler to design and operate. Users initiate the contacts by dialing a telephone number and entering a password that identifies them to the system, which then confirms their name and accesses stored data. All of our early systems were inbound systems. Outbound systems have to account for the increasingly complex set of possibilities of whom or what answers the phone. The details of this are beyond the scope of this paper, but a difficult challenge is designing the systems so it reliably can differentiate between a live person and an answering machine or voice mail system answering the call. In addition, protocols have to be specified to handle a number of situations including scheduling call backs if the line is busy, not answered, if the user is not home or home but not available to take the call. TLC, for example, will leave a short message on an answering machine, but although it might call back repeatedly, it will only leave one message per day and no more than a few per week. TLC also produces daily reports on calls completed, calls partially completed, and the results of other call attempts to be reviewed by project staff. Other protocols specify when a staff person is to contact the user by phone or mail to trouble shoot connection problems.

Beyond technological and practical issues, there are pros and cons of each method of initiating a TLC conversation. An inbound system gives the user more control over the initiation of an interaction; however, we have found that inbound systems result in fewer than prescribed contacts. Out-bound systems result in better adherence to the planned contact schedule[39]. Our experience tells us that outbound systems are superior. Additionally, they have the capability to accept inbound calls from the user which preserves some user control. We have also given users the option to switch from outbound to inbound calls.

Automated Interactions and Response Options. A design choice that has become available recently is whether to use automated speech recognition (ASR) technology. In ASR systems, the computer is "listening" for a finite set of responses after each question. Users are provided with this set of responses and can speak their answer into the telephone. Although many times, a user may reply with an answer that is different than the set of responses provided. For example, after a Yes/No

question, the system says, "*Please say yes or no*", but the user might answer "*yes*" or "*yeah*." ASR technology can be programmed to accept common synonyms (e.g., yea, yup, nah, nope, etc). If the system does not understand a reply, there is a protocol for querying the user. ASR can also understand multiple word answers. In one of our TLC diet programs, we provide users with a set of five benefits of a healthy diet and ask the user to indicate their choice, e.g., the system says: *"If you are interested in hearing about how diet can improve your skin, say 'skin'. If you're interested in how a healthy diet can lower your blood pressure, say 'blood pressure'."* If no answer has been received by the end of the list, the system will say, "*Please say skin, blood pressure, . . . , or none now.*" The ASR technology we use allows the caller to 'barge in,' meaning the system is listening for the user's answer while the question is being asked or before the set of possible responses are being listed.

There are definite advantages and disadvantage of using ASR technology. One advantage of ASR is that the call functions more like a human conversation. Given that many phones have keypads on the handset, ASR also eliminates the need to move the phone away from his or her ear. The disadvantages of ASR are that it has some problems in recognizing certain spoken responses such as large numbers (e.g., blood pressure readings). We have found that there is a small subset of users who get frustrated enough with the system not understanding their answers that they stop using the system. Another small subset of users has trouble keeping the system's technological limitations in mind—a phenomenon that might be exacerbated with the use of ASR systems that function more like a human conversation by "listening" and "understanding" responses. This more human-like conversation may increase the users' expectations of the system and frustration with its limitations. Although our recent systems have incorporated ASR technology; however, we are uncertain whether ASR represents a clear improvement over touch-tone technology given its current limitations.

## 5. Steps to writing a comprehensive dialogue

We have described the background and parameters of how to structure a dialogue system, as well as important specific design considerations. The next section will describe how we approach the writing of a TLC dialogue system as well as the overall qualities we feel are important for a high quality script. The script for the dialogue is generally written by either an expert in the content area or an expert assisted by a professional writer. The author (script writer) needs to have a good understanding of the design specifications (e.g., the goals of the system, the theory to be applied, duration and schedule of contacts) that have been discussed in the previous sections of this paper.

### 5.1 Characteristics of the optimal dialogue systems

It is important that the user experiences the dialogue system as an engaging and valuable conversation. This enhances the user's attentiveness and increases the likelihood that the person will respond positively to the behavioral counseling provided. It also contributes to the user's adherence to the recommended call schedule, which is important for programs that entail multiple contacts over time. Creating natural sounding and engaging conversations using automated system, given the present technological limitations (e.g. unable to comprehend free speech) takes a concerted effort.

Based on our extensive experience in writing and evaluating automated dialogue intervention systems, we believe effective systems are based on dialogue that has several important qualities including that the dialogue: (1) is optimized for spoken communication with lay people, (2) endows the system for human-like characteristics to resemble real conversations, (3) is personalized to the individual user, (4) maximizes interactivity, (5) balances repetition and novelty of content, (6) mixes system and user control of the conversation, and (7) maximizes pseudo-intelligence.

Writing messages that are suited for communications that are heard by a lay population is often a surprisingly difficult task for health professionals without previous experience in dialogue construction. Most health professionals, especially those who are researchers and academics,

principally write research grant proposals and journal articles that are addressed to a professional audience who will read rather than hear what is written. It is a two fold adjustment to both write for a much lower educational level and to write conversation instead of text to be read. What often works best is to pair a content expert with a professional writer who is known to be good at writing this type of dialogue. We also recommend that writers read out loud the dialogue they have written, even if only to themselves. One important point to remember is that only a limited amount of material can be kept in mind by the listener at any one time. Complex lines of reasoning or long lists might work in a text based intervention, but are generally not appropriate for spoken communications.

Our research suggests that our systems are quite successful in emulating real conversations as users clearly anthropomorphize the TLC voice [40]. We deliberately invest the voice with human qualities in order to better engage the user. First, we use actors and actresses to record the scripts. This ensures that the voice is easy to understand and is rich with variation in inflection and tone (although computer generated voices are improving in these qualities). Second, conversation structure and content is designed to emulate the characteristics of a human conversation. For our programs which are focused on care of a medical condition, the TLC dialogues are usually emulating a health professional speaking with a patient or consumer over the telephone. We have conducted a number of observational studies in which we record the interaction between training health professionals, most often registered nurses and nurse practitioners, with patients on the phone in order to identify the characteristics of human health professional delivery over the telephone. From what we have learned, we have designed the content of the TLC conversations, including the questions and declarative statements, the order of presentation of content, how the system responds to questions and the words, sentence structure and tone used, to closely match the user's expectations of what a health professional (e.g., a nurse) might ask, respond and sound like. In-depth interview studies have demonstrated that users believe that the TLC scripts successfully emulate a human health professional [17].

For programs that are less clinical in nature, such as diet improvement programs, the script is less formal, and we consider the voice to be characterized more as a friendly but expert advisor or coach. As if one was writing the dialogue for a character in a play, it is best to have a clear sense of the role of the dialogue voice of the intervention. To have a consistent presentation it is also best to only have one or two 'authors' of the script (although it can have many designers) with one doing the final editing. In programs that target multiple behaviors and are written by separate teams of content experts and writers, we use a different voice, which we give a first name, for each behavior. This allows each of the voices to take on the personality conveyed by the individual script writer and for the user to identify it with one behavior.

Other ways we 'humanize' the automated dialogue is to include some occasional humor into the script, have the system check-in with users about their status within the conversation, (e.g. "*Ready to go on?"*), or have it occasionally be self-deprecating ("*I know I am only a computer, but ...*"). The use of humor has to be judicious as to not undercut the seriousness of the topic or the expertise of the TLC 'advisor' and is not appropriate when dealing with sensitive health topics.

Although users might anthropomorphize the voice, most understand that the system is a computer but "go along" with the fiction that they are interacting with a human health professional [10]. In a similar way that movie goers 'suspend disbelief', users, in order to maximize their participation and enjoyment, set aside their understanding of what is going on and interact with TLC as if it were a person. In qualitative interviews, we usually hear that users 'feel' that they are interacting with a person, even though they 'know' it is a computer system.

Further anthropomorphizing of TLC is engendered by maximizing the system's pseudo-intelligence. Of course, the system is not truly intelligent; however, we wish it to function in a manner that as closely as possible emulates how an excellent clinician would function. Evaluations of TLC systems indicate that the users perceive it in that manner. However, they also identify "lapses in

intelligence." For example, several users who used a TLC physical activity application complained that sometimes they were not able to exercise due to the inclement weather but the closed-ended nature of the conversation made it impossible to communicate that fact to TLC. For this reason, it is important that users of dialogue systems be educated about the limits of its "intelligence" such as its inability to understand very particular circumstances the person might be in. Likewise, it is important for the designers to identify particular instances in which a section of script does not allow an option that some subset of users might expect. This recognition can only occur if systems are thoroughly tested before release, including testing by some members of the target population, and their performance closely monitored after release. In the meantime, to reduce user frustration, in some TLC applications we allow users to leave a message for the study staff at the end of each conversation in which they can tell us what they could not tell the system.

The hallmark of advanced dialogue systems for health behavior change is the ability to have the communications be very specific to the user. Using terminology codified by Krueter [41, 42], this is done in three ways: personalizing, tailoring, and individualizing, each described below.

First we use information about the user to 'personalize' the conversation. The data used for this is not necessarily relevant to the user's health status or to specific behavior change strategies, but makes it evident that the system knows who it is talking to. This includes referencing the user's name, knowledge of the user's family members, names of caregivers or the neighborhood they live in. Generally, this helps build a sense of a relationship with the user. It is worth noting that one can overuse this feature. In early TLC programs, we used the person's preferred name very frequently, such as "*as you know, John, your blood pressure of 160/100 is higher than what your doctor would like it to be.*" Overuse of the person's name can be grating.

Second we 'tailor' to the subgroup or subgroups the user belongs to. This might be the subgroup that the whole user population belongs to, as when we have tailored an intervention to a racial or ethnic group, or there might be a number of subgroups within a population, such as different age groups, genders, or groups defined by the TTM stages of change described above that dialogue content is tailored to.

Most importantly, we 'individualize' the content of the messages based on specific information we collect about the users. This information can either be collected before the beginning of the intervention or during the TLC calls. Although it is often more efficient to collect data before the intervention either from surveys or medical records, information that is collected during the call and used immediately promotes a more open and conversational experience. It is also important to collect current information on variables that might change over time. The TLC system stores this data and it can be used over time to communicate patterns of change. Studies of TLC users have indicated very positive responses when TLC "remembers" something about the person [43]. Individualization on variables important to either the user's clinical status or to the process of behavior change allows conversations to be relevant, efficient, theory-based, and effective.

There are, however, clearly limits to the scalability of this process when using finite state machine architecture. If one is using multiple variables to individualize messages, the number of messages that have to be written, recorded, and programmed can quickly become very large. For example, if different versions of a message are being written based on the 5 stages of change, and three levels of self-efficacy (e.g., low, medium, and high), then 15 messages would have to be written. If, subsequently, it is decided to add in whether the user's self-efficacy has increased, decreased, or stayed the same since the last call, there are now 45 messages to be created, yet, because the user will only hear one of these during any one call, this might only account for 15 seconds of dialogue. For this reason it is important to be clear on what are the most important variables. The theoretical basis of the system should provide a guide to this prioritizing of dimensions of individualization.

In TLC conversations we carefully mix repetition of call flow and content with new and even unexpected material. Repetition serves several purposes. For example, since in nearly all TLC programs the users are being assessed regularly, often about the same phenomena (e.g. amount of physical activity, etc), it makes the conversation go faster and more smoothly if the assessment appears in the same part of the conversation and is done in the same way each call. We believe that such repetition of sections provide almost a comforting familiarity for some users. Novelty is also necessary and comes from introducing new topics or new approaches to the same issue and in reviewing changes in user's behaviors and/ or attitudes over time.

As mentioned, TLC systems are designed as finite state machines in which the flow of the conversation, though complex is completely pre-determined. This ensures that the objectives of the program during each contact are met and that the interview is complete and appropriate and the counseling is consistent with the intervention plan. At the same time, we attempt to introduce multiple opportunities throughout the TLC conversations for the user to direct the conversation in ways that make sense to that person. This includes asking the user to choose between topics to 'talk' about during a call or to decide whether to hear more detailed information about a topic or to have content they have already heard repeated. In other situations, a decision might involve input from both the user and the system. For example, in a diet program for hypertensives that focuses on consumption across four food groups, an initial decision has to be made about which food group to start with. In this program, the system allows the user to know which food group it would recommend based on the user's present consumption. Then, the user is asked to decide what food group to begin with. A similar interaction is often used in setting weekly goals, with the system either making recommendations first or having the user lead with a new goal and the system suggesting adjustment if it seems either too low or too high to be optimal. Whenever it is reasonable we design systems so the user's choice prevails.

5.2 Flow diagrams

One common challenge in writing dialogue systems for behavioral change interventions is how to manage the complexity of the potential conversations. This can be managed to some degree when flow diagrams are used. Our laboratory uses a commercially available software package designed for creating flow diagrams. A series of flows are typically written at progressing levels of detail from general to specific. We begin with what is called a "top-line" flow which outlines the more general objectives of the overall script and the basic topics addressed in each call. The theory chosen to guide the intervention should be evident in the flow's top-line. Figure 1 displays a top-line flow for the fruit and vegetable module of a diet intervention. Notice that it is general and incorporates the theoretical constructs to be targeted.

[Insert Figure 1 about here]

As the script is outlined, the top-line flow is fleshed out with more specifics for each call, e.g., introduction, assessment, feedback, education, goal setting, and closing statement. This process of diagramming the calls continues until there is enough detail to suit the writer's needs. Figure 2 shows the more specific flow of the call that was represented in the box labeled 'Confidence Rating and Feedback' in Figure 1. We have found that more detailed flows can make it easier to write the script. As illustrated in Figure 2, the flow has a branching structure to describe the various paths that a conversation can take and there is a phrase in each box conveying the message content to be conveyed to the user.

[Insert Figure 2 about here]

5.3 Collaborating with the technical staff

The system design team includes more than the experts in the content area, it also includes the programming staff, system testers, and often a liaison between the writers and the programmers. Our laboratory uses a liaison person to review and critique dialogue scripts, clarify logic, optimize conversational dialogue, and put the script into a format that enables it to be computer programmed e.g., includes 'go to' and 'variable' statements. Technical comment statements can be added by the authors, the liaison or the programmers to document the logic of a section of script. To this end, the flow diagram developed by the authors is used by the programmers to understand how the script flows.

There are many aspects of the script that are repetitive such as verifying what a caller has vocalized. More specifically, a typical script contains many "yes/no questions" with the question followed by the statement "*say yes or no now*" If yes, go to line 4, If no, go to line 5. (Note that while the early textual forms of the script use branch logic for simplicity, this is converted into finite state machine logic and runtime software implemented using state-of-the-art structured programming techniques.) The liaison can complete these dialogue components for the author. Another role of the liaison is determining whether all branching logic possibilities have been accounted for. Often there are responses to a question, which since they are atypical, are not considered by the script writer, but the control system must account for. Related to this, the caller may provide improbable responses, i.e., fall outside of the range of logically possible or acceptable answers. TLC may ask, "*Please tell me the number of days that you exercised during the past week?*" If the caller answers 9, then TLC will say, "*I am sorry, I heard 9. I was asking about the number of days you did moderate activities in the last week. The number cannot be higher than 7. Please say the number of days again. Say zero if you do not do any moderate physical activity. You may also enter the number of days using your key pad, entering a number from zero to seven."* The liaison will add these types of responses throughout the script.

5.4 Suggestions for Dealing with common Dialogue Situations

From our extensive experience writing dialogue for TLC systems we have developed approaches to some common situations that arise when trying to translate a behavioral change intervention into an automated dialogue system. These approaches range from simple to more complex. One situation is when it is necessary to have the user leave the phone temporarily either to get something (pen and paper) or to perform a task (weigh themselves). In these instances, we instruct the user to either press one or say "*ready*" when they return to the phone. We have learned that sometimes users forget this instruction when they return to the phone so we have the system repeat this instruction every 5 seconds so the user is likely to hear it when they pick up the phone. Another aspect of this situation is how long to have the line stay open because occasionally the user doesn't return to the phone. Since this uses some system resources (mainly a telephone line), the level of system resources somewhat determines how long the line can be left open. We usually use a 5 minute timeout for simple tasks such as getting paper and pencil to write down something (e.g., a phone number, homework assignment), and have found that a longer time such as 10 minutes is needed if subjects have to accomplish a task (e.g., taking their blood pressure or finding something that might be misplaced, like a report we have sent them.)

Another common situation is when there is extensive information to deliver to the user. We have found that lengthy monologues go against the nature of a good conversation and make the experience less engaging. The amount of information one might easily put into a paragraph in a written communication is often too long for a phone system if not broken into smaller segments. In instances in which there is extensive material to present on a single topic, we attempt to break the

material into smaller segments of only a few sentences each, usually with questions between the segments. These questions might merely serve to keep connected with the users, such as "*did you get that?*" or a rhetorical question to break up the content. Utilizing quiz type questions is another way to keep content conversational. Then you can provide detail on the topic regardless if the user answers correctly or incorrectly. For example, if the system is trying to present benefits of exercise, the system could say: "*And did you know that regular exercise will help you sleep better and think clearer? Say Yes or No now.*" The next message provides a bit of detail on this, but begins with: "*that's right*!" if they responded yes, and "*Well, it does*" if they responded no. Another way to increase interaction and decrease length of utterances is to present information at multiple levels of detail. For example, in a TLC health information program [7], we developed a program of over 50 health information topics derived from the Harvard health letter, a health newsletter published by Harvard University Medical School publications. To deal with the large amount of factual information, we broke the content up into sections and in each section provided an brief overview of the content in that section. We then asked if the user wished to learn more about this topic, and if they did, provided more detail before moving on to the next sub-topic. When a lot of information has been communicated, it is important to include brief reviews of information along the way.

Asking certain types of questions can be more difficult in dialogue systems than text based systems. For individuals of average or higher literacy, a printed question can be read and responded to faster than listened to and answered. Scale instrument with many items can be too burdensome for a dialogue system, as well as being distinctly non-conversational. The other type of question that is difficult to construct is when the purpose is to pick between many different options. Since the user cannot scan the list, they are limited by how many choices they can keep in mind at any one time. We found that if there are four to five choices, they can all be presented and the user can be asked to pick the one that they want to hear about most or first. If the list is a set of topics and we want the user to hear detail on all or most of them during the program, a common approach is to present the same list during each subsequent call, minus the options already taken. When we have a question with many answers, which is often the case in these systems, a multi-step approach has to be taken, and these steps can be designed in a number of ways. For example, a common section of our scripts is to give feedback on how to overcome barriers to a behavior, where there are often 10 to 20 possible barriers. Optimally the user would be able to choose the barrier that they would gain the most from hearing, but this is difficult to do with this many choices. Barriers can be presented one at a time, and users are asked to state whether the barriers are important to them or not. Then only the endorsed items are repeated and they choose from the smaller set of items. This can be time consuming, and depending on the behavior, users can endorse most of the original set, inadvertently perpetuating the problem. A second option is to divide the list into subgroups and ask them to pick one barrier of the first group (or say '*none*') and then give feedback on that one. Then the system can move onto the next subgroup either during the same call or in subsequent calls. Combinations of these approaches can be designed. For instance, one could take the first approach for a subset of items so the most important barrier of the first subgroup is addressed, and then do the same in the second subgroup. Exactly which method is optimal depends on how much contact time is available both during a call and across calls for the topic, how many items are likely to be relevant to the user, and how important it is to give them feedback on only the most relevant items.

## 6. Conclusion

In this paper we have endeavored to provide a road map to the essential design considerations, component tasks, and content considerations necessary to create effective, telephone-based, automated dialogue interventions. These recommendations are based on our experience developing and

evaluating these systems over the past two decades. The power of these systems is likely due to the combination of three components: new and improving computer technologies, better understanding of how to assist individuals to improve health-related lifestyle behaviors, and the use of the telephone, the oldest, widely-used telecommunication technology, to engage in 'human' conversations. This combination allows for a multitude of potential applications that have only started to be explored.

Interest in the use of automated dialogue systems to improve health related outcomes both within and outside of the health care delivery system is increasing rapidly as its potential to improve both individual and public health outcomes at a relatively low cost is being recognized. As this paper was focused on the writing of dialogue, we did not present the use of telephony for disease management as those systems involve a number of other functions such as monitoring of patient health conditions and alert generation, but as the cost of health care continues to rise the use of telephony in disease management is likely to also increase.

There are other potentials that are only beginning to be explored. There are possibilities to integrate such systems with other forms of automated health interventions (e.g., web-based interventions) as well as with the increasingly computer-based health care delivery system (e.g., EMR's). An area of research we are investigating is whether different people prefer and get more benefits from interventions depending on the type of communication delivery channels used (e.g. phone, web, or print) which could inform the matching of communication channel at the individual level. Improved computer generated voices and text to speech functionality will allow a more streamlined development process - dialogue will be heard and tested virtually as it is created. Improvements in the quality of computer generated speech could reduce the cost of system development and increase system flexibility.

As mentioned, there are practical limits to the scalability of finite state machines on the number of factors that can be simultaneously considered in tailoring content. More advanced computational approaches (see T. Bickmore, T. Giorgino, 2006)[44] such as dialogue planning techniques have the potential to overcome these limitations. As they are still in the early stages of development, it is thus far unclear whether use of these approaches will achieve a level of effective tailoring that can be obtained with meticulously constructed dialogues delivered using finite state machine architecture. This is an important researchable question that investigators need to address.

The proven effectiveness of TLC systems notwithstanding, there are also a number of challenges ahead. There is much to understand about how best to design and implement these systems. Despite overall effectiveness, there are still many users who do not benefit from these systems or only benefit minimally. And lastly, these systems have largely been developed and tested in the context of federally funded research projects, and it is time to move to the phase of dissemination and evaluation of these low cost interventions in real world situations.

[1] Delichatsios HK, Friedman RH, Glanz K, Tennstedt S, Smigelski C, Pinto BM, et al. Randomized trial of a "talking computer" to improve adults' eating habits. Am J Health Promot 2001;15(4):215-24.
[2] Glanz K, Shigaki D, Farzanfar R, Pinto B, Kaplan B, Friedman R. Participant reactions to a computerized telephone system for nutrition and exercise counseling. Patient Educ Counseling 2003;49(2):157-163.
[3] Migneault JP, Friedman RH, Resnicow K, Campbell M, J. D-J, Hill M, et al. Design of an automated multi-behavioral intervention for hypertensive African-Americans. Annals of Behavioral Medicine 2003(Supplement):S161.
[4] Jarvis KL, Friedman RH, Heeren T, Cullinane PM. Older women and physical activity: using the telephone to walk. Womens Health Issues 1997;7(1):24-9.
[5] Pinto BM, Friedman R, Marcus BH, Kelley H, Tennstedt S, Gillman MW. Effects of a computer-based, telephone-counseling system on physical activity. American Journal of Preventive Medicine 2002;23(2):113-20.
[6] King A, Friedman R, Marcus B, Napolitano M, Castro C, Forsyth L. Increasing regular physical activity via humans or automated technology: The CHAT trial. Annals of Behavioral Medicine 2003(Supplement):S111.
[7] Ramelson H, Bassey B, Friedman R. The use of computer telephony to provide interactive health information. In: American Medical Informatics Association; 2003; 539-543.
[8] Friedman RH, Kazis LE, Jette A, Smith MB, Stollerman J, Torgerson J, et al. A telecommunications system for monitoring and counseling patients with hypertension: impact on medication adherence and blood pressure control. American Journal of Hypertension 1996;9(4 Pt 1):285-92.
[9] Young M, Sparrow D, Gottlieb D, Selim A, Friedman R. A telephone-linked computer system for COPD care. Chest 2001;119(5):1565-75.
[10] Farzanfar R, Finkelstein J, Friedman R. Testing the usability of two automated home-based patient-management systems. Journal of Medical Systems 2004;28:143-153.
[11] Migneault JP, Rakowski W, Friedman RH, Marks L, Clark M. User engagement with an automated telephone intervention to improve on-time repeat mammography. Annals of Behavioral Medicine 2005;D33(Supplement):S85.
[12] Morris M, Sparrow D, Gottlieb D, Selim A. A telephone linked computer system for COPD care. Chest 2001.
[13] Young DR, Gittelsohn J, Charleston J, Felix-Aaron K, Appel LJ. Motivations for exercise and weight loss among African-American women: focus group results and their contribution towards program development. Ethnicity & Health 2001;6(3-4):227-45.
[14] Adams W, Fuhlbrigge A, Miller C, Panek C, Gi Y, Loane K, et al. TLC-asthma: an integrated information system for patient-centered monitoring, case management, and point-of-care decision support. Proc AMIA Symp 2003:1-5.
[15] Cohen J. Statistical power for the behavioral sciences. 2nd ed. Hilldale, NJ: Erlbaum; 1988.
[16] King A, Friedman R, Marcus B, Napolitano M, Castro C, Forsyth L. Increasing regular physical activity via humans or automated technology: 12-month results of the CHAT trial. Annals of Behavioral Medicine 2004;27(Supplement):S044.
[17] Kaplan B, Farzanfar R, Friedman R. Personal relationship with an intelligent interactive telephone health behavior advisor system: a multi-method study using surveys and ethnographic interviews. International Journal of Medical Informatics 2003;71:33-41.

[18] Farzanfar R, Frishkopf S, Friedman R, Ludena K, Ward A. Humanizing the machine: refinement of an automated patient-management system. Annals of Behavioral Medicine 2004;27(Supplement):S154.
[19] Paperny DM. Computerized health assessment and education for adolescent HIV and STD prevention in health care settings and schools. Health Education and Behavior 1997;24(1):54-70.
[20] Mears M, Coonrod DV, Bay RC, Mills TE, Watkins MC. Routine history as compared to audio computer-assisted self-interview for prenatal care history taking. Journal of Reproductive Medicine 2005;50(9):701-706.
[21] Pate RR, Pratt M, Blair SN, Haskell WL, Macera CA, Bouchard C, et al. Physical activity and public health. A recommendation from the Centers for Disease Control and Prevention and the American College of Sports Medicine. JAMA 1995;273(5):402-7.
[22] Resnicow K, Braithwaite R, Kuo J. Interpersonal intervention for African American adolescents. In: Wilson D, Rodrigue J, Taylor W, editors. Adolescent health promotion in minority populations. Washington, D.C.: American Psychological Association; 1997. p. 201-228.
[23] Hardy ME. Perspectives on science. In: Role theory: perspectives for health professionals. 2nd ed. Norwalk, CT: Appleton & Lange; 1988.
[24] Kerlinger FN, Lee HB. Foundations of behavioral research. 4th ed. Orlando, Fl: Harcourt College Publishers; 2000.
[25] Bandura A. Social foundations of thought and action: a social cognitive theory. Englewood Cliffs, N.J: Prentice Hall; 1986.
[26] Bandura A. Self-efficacy: the exercise of control. New York: W.H. Newman and Company.; 1997.
[27] Becker MH. The health belief model and personal health behavior. Health Education Monographs 1974;2:324-508.
[28] Fishbein M, Ajzen I. Belief, attitude, intention and behavior: an introduction to theory and research. Boston, MA: Addison-Wesley; 1975.
[29] Kanfer FH. Self-regulation and behavior. In: Neuringir C, Michael JL, editors. Behavior modification in clinical psychology. New York: Appleton-Century Crofts; 1970.
[30] Triandis HC. The analysis of subjective culture. New York: John Wiley and Sons; 1972.
[31] Bandura A. Self-efficacy: toward a unifying theory of behavior change. Psychological Review 1977;84:191-215.
[32] Prochaska J, Velicer W, Fava J, Rossi J, Tsoh J. Evaluating population-based recruitment approach and a stage-based expert system intervention for smoking cessation. Addictive Behaviors 2001;26:583-602.
[33] Prochaska J, Velicer W, Rossi J, Redding C, Greene G, Rossi S, et al. Multiple risk expert systems interventions: impact of simultaneous stage-matched expert system interventions for smoking, high-fat diet, and sun exposure in a population of parents. Health Psychology 2004;23(5):503-16.
[34] Velicer W, Prochaska J, Fava J, Laforge R, Rossi J. Interactive versus noninteractive interventions and dose-response relationships for stage-matched smoking cessation programs in a managed care setting. Health Psychology 1999;18(1):21-28.
[35] Rogers R, Prentice-Dunn S. Protection motivation theory. In: Gochman D, editor. Handbook of health behavior research: Vol. 1. Determinants of health behavior: Personal and social. New York: Plenum; 1997. p. 113-132.
[36] Weinstein ND. The precaution adoption process. Health Psychology 1988;7:355-386.
[37] Prochaska J, Velicer W. The transtheoretical model of behavior change. American Journal of Health Promotion 1997;12:38-48.
[38] Prochaska JO, DiClemente CC. Transtheoretical therapy: toward a more integrative model of change. Psychotherapy theory, research, and practice 1982;19:276-288.

[39] Wright J. Utilization of automated telephone systems for diet and physical activity: a cross study analysis. Annals of Behavioral Medicine 2005;D33(Supplement):S84.
[40] Kaplan B, Farzanfar R, Friedman R. Ethnographic interviews to elicit patients' reactions to an intelligent interactive telephone health behavior advisor system. In: Lorenzi NM, editor. American Medical Informatics Association, Annual Symposium; 1999; Bethesda, MD; 1999. p. 555-559.
[41] Kreuter M, Lezin N, Yung L. Evaluating community-based collaborative mechanisms: Implications for practitioners. Health Promotion Practice 2000;1:49-63.
[42] Kreuter M, Farrell D, Olevitch L, Brennan L. Tailoring health messages: customizing communication with computer technology. Mahwah, NJ: Lawrence Erlbaum Associates; 2000.
[43] Farzanfar R, Frishkopf S, Migneault J, Friedman R. Telephone-Linked Care for Physical Activity (TLC-PA): a qualitative evaluation of the use patterns of an information technology program for patients. Journal of Biomedical Informatics 2005;38:220-228.
[44] Bickmore T, Giorgino T. (To be determined). Journal of Biomedical Informatics 2006;This issue.

**Figure 1**

**Click here to download high resolution image**

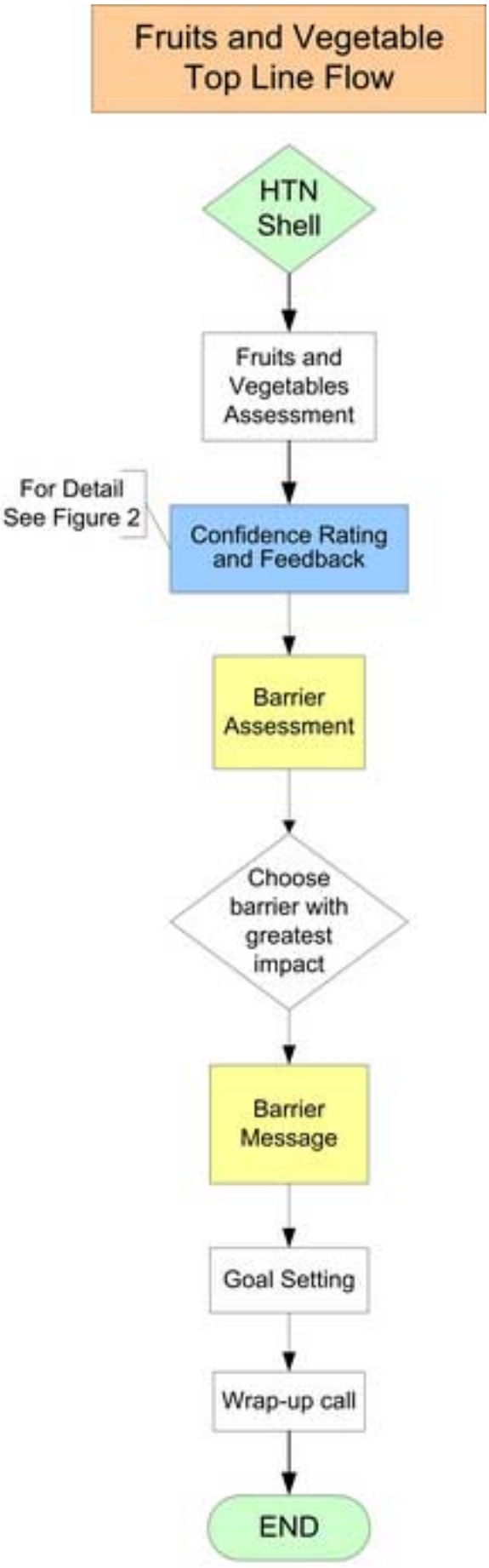


Fig. 1

**Figure 2**

**Click here to download high resolution image**

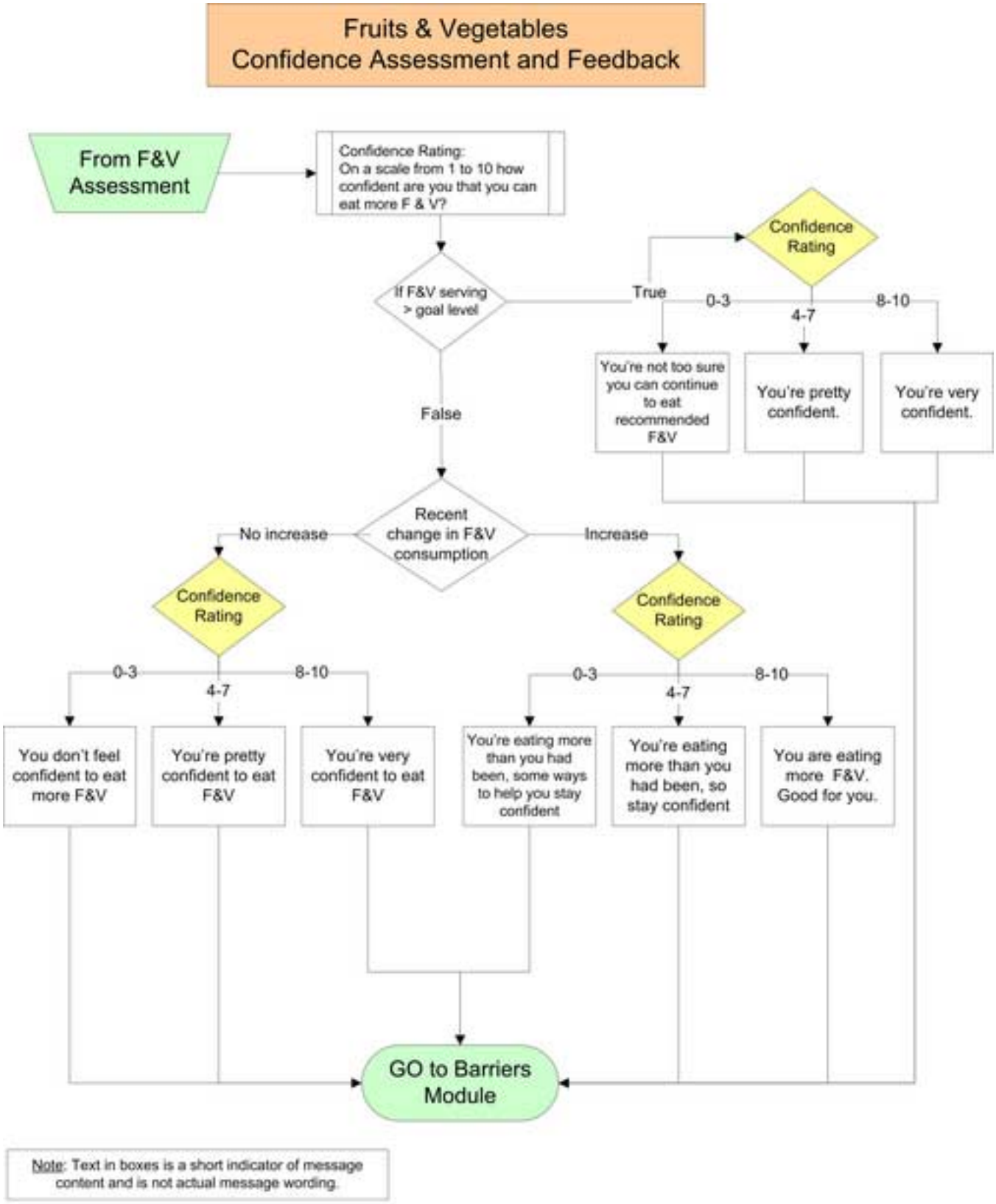


Fig. 2