A Reusable Framework for Health Counseling Dialogue Systems based on a Behavioral Medicine Ontology

Timothy W. Bickmore

Assistant Professor, College of Computer and Information Science

Northeastern University

Boston, Massachusetts

Daniel Schulman

College of Computer and Information Science

Northeastern University

Boston, Massachusetts

Candace L. Sidner

Research Professor, Department of Computer Science

Worcester Polytechnic Institute

Worcester, Massachusetts

Under Review

Abstract

Automated approaches to promoting health behavior change, such as exercise, diet, and medication adherence promotion, have the potential for significant positive impact on society. We describe a theory-driven computational model of dialogue that simulates a human health counselor who is helping his or her patients to change via a series of conversations over time. Applications built using this model can be used to change the health behavior of patients and consumers at low cost over a wide range of media including the web and the phone. The model is implemented using an OWL ontology of health behavior change concepts and a public standard task modeling language (ANSI/CEA-2018). We demonstrate the power of modeling dialogue using an ontology and task model by showing how an exercise promotion system developed in the framework was re-purposed for diet promotion with 98% reuse of the abstract models. Evaluations of these two systems are presented, demonstrating high levels of fidelity to best practices in health behavior change counseling.

1. Introduction

Poor health behavior, such as lack of physical activity and unhealthy dietary habits, are among the leading causes of death and chronic disease in the United States [1], yet a large proportion of Americans fail to meet guidelines published by U.S. public health authorities. For example, participation in moderate amounts of physical activity has important health benefits, including beneficial effects on risk factors for disease, disability, and mortality [2-6], yet a substantial proportion of the U.S. adult population remains underactive or sedentary relative to the current CDC recommendation of 150 minutes per week of moderate intensity aerobic activity for adults [2]. Fruit and vegetable consumption plays a protective role in a large number of epithelial cancers [7, 8], and is associated with reduced risk for heart disease, stroke, and hypertension [9-15], yet only 23% of American adults meet the (year 2000) recommendation to consume five or more servings of fruit and vegetables per day [16, 17].

The "gold standard" for health behavior change remains one-on-one, face-to-face interaction with an expert human counselor, who has at his or her disposal a range of theory-based interventions that can be deployed based on the needs of the patient (e.g., in response to a question or recent experience) and the conversational context (e.g., building upon information disclosed earlier in the conversation or using a patient question to lead into an anecdote) [18]. Unfortunately, one-on-one interventions only have the potential to reach a very small portion of the population, resulting in very low impact (efficacy multiplied by recruitment rate[19]). To reach a larger audience, a wide range of media-based and automated health behavior change interventions have been developed over the last two decades. Of these, the ones that come closest to the "gold standard" are those which autonomously communicate with patients using natural language dialogue, emulating the behavior of an expert counselor as closely as possible (The appendix provides sample transcripts of such dialogues for exercise promotion.)

Many health counseling dialogue systems have now been developed and evaluated in randomized trials. The design of these systems is typically based on one or more theoretical frameworks or counseling methods from behavioral medicine, such as the Transtheoretical Model (also known as the "stages of change" model) [20], Social Cognitive Theory [21], and Motivational Interviewing [22]. They have been designed to promote a range of health behaviors including diet, physical activity, smoking cessation, medication adherence, and chronic disease self-care regimens, and most have been shown to be effective on at least one outcome measure, compared to standard-of-care or non-intervention control conditions [18, 23-27].

1.1. The Need for Reusable Computerized Behavioral Interventions

Although many health counseling dialogue systems have been shown to be efficacious, they are still being developed with little concern for reusability. In addition to representing a waste of money and resources due to duplications of effort by researchers, these issues impact the ability of these systems to be disseminated to health care practitioners. Adapting a computerized behavioral intervention developed for a research study for large-scale deployment by an institution or public health department typically requires significant modifications to the original

system, and these modifications often represent an insurmountable barrier to dissemination [28]. These issues have recently become a concern of funding agencies who have been supporting the development of computerized behavioral interventions and are now interested in seeing these systems disseminated [29, 30].

1.2. How Ontologies Support Extensibility, Reuse, and Dissemination

An ontology is a taxonomic description of the concepts in an application domain and the relationships among them [31]. Computational ontologies have three primary uses in knowledge-intensive software systems: (1) they provide a formalism that can facilitate clarification and description of the fundamental concepts in a domain through consensus of experts; (2) they can facilitate interchange of information among diverse systems by describing, at various levels of detail, the kinds of data entities that can be exchanged, independent of the particular names the entities are given in each system; and (3) they can promote reuse of software components through the development of ontologies of the components themselves (descriptions of what the components do) and the data entities they operate on.

Ontologies can be used in several ways in a computerized health behavior change intervention to promote reuse and interoperability. For example, many kinds of knowledge can be re-used across interventions, such as which constructs are important to assess for a given health behavior change theory, which dialogue strategies can be used to assess the constructs, and the specific language that should be used to talk about a given health behavior. Knowledge about how to structure an intervention counseling session can also be abstracted and re-used through hierarchically decomposable task models, specifying, for example, that in general, a session is opened with a greeting followed by small talk, and is closed with a farewell.
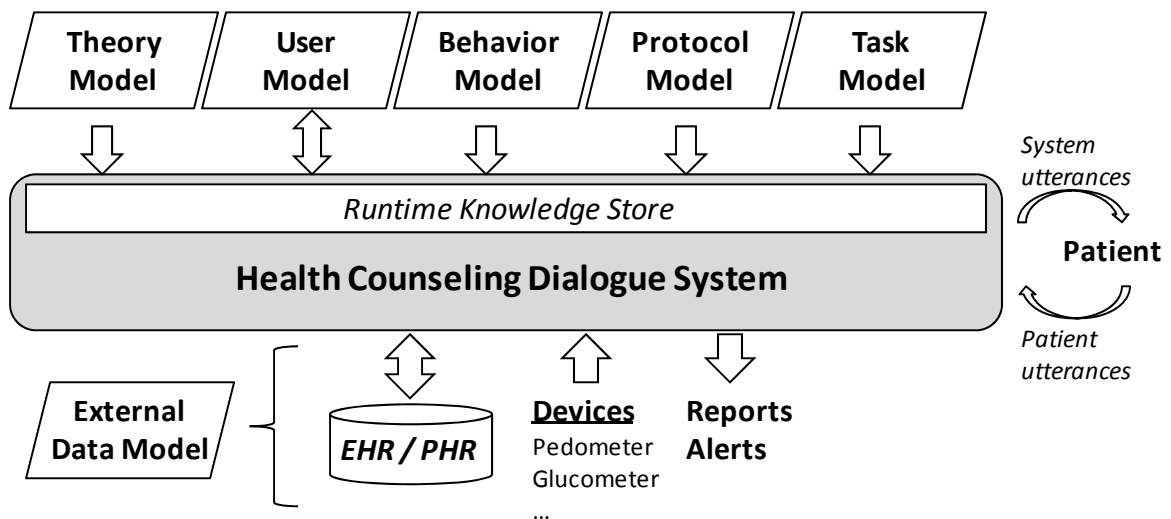

Figure 1. Reusable Knowledge in Health Counseling Dialogue Systems

The kinds of knowledge that can be abstracted and reused are outlined in Figure 1:

A ***Theory Model*** contains knowledge of the behavioral medicine theory the intervntion is based on, including relationships among health behavior change theories, constructs, and counseling techniques (as in [32]). This knowledge can be readily reused across all computerized interventions that are based on the theories represented.

A ***Behavior Model*** contains knowledge of how health behavior change theories are applied to a specific health behavior, such as exercise promotion or smoking cessation. This includes general knowledge, such as whether the intervention is acquisition or cessation oriented, and whether the behavior can be shaped (incrementally modified) or not. This can also include very specific information such as homework exercises or tips that can be recommended to patients, or utterances or phrases the dialogue system uses in its conversations (e.g., that the behavior is called "walking", that the units of measure are called "steps", etc).

A ***Protocol Model*** contains knowledge about a particular behavior change intervention, including fixed parameters such as duration of the intervention and expected number of patient contacts per week, and ultimate behavioral criteria such as achieving "30 minutes a day of moderate physical activity on most days of the week" [33-35].

A ***User Model*** contains knowledge about a particular patient, including fixed "tailoring" parameters that affect the intervention and messages (e.g., gender, age, ethnicity) [36], information about the patient's medical condition (e.g., from an EHR or PHR), construct measures (e.g., self-efficacy, stage of change), and information that is dynamically updated over the course of an intervention (e.g., daily behavioral goals).

An ***External Data Model*** describes data inputs and outputs from the system. This includes data that can be used by the system, such as medical data from an EHR or PHR, or measurement data from devices such as pedometers (for exercise promotion) or glucometers (for diabetes self-care promotion). This also includes data that can be output by the system for use by clinicians, including measurements and patient status information to the patient's EHR or PHR, detailed transcripts or summary reports of individual counseling sessions (e.g., a psychiatrist might want to know *exactly* what the patient said in a given session), and alerts describing emergent conditions to be broadcast in a timely manner to appropriate clinicians. Describing these inputs and outputs in an ontology promotes interoperability and reuse.

Finally, a ***Task Model*** contains knowledge about *how* a health behavior change counseling system enacts the intervention, taking all of the above information into account. While the other knowledge representations are conceptual in nature, knowledge in the Task Model is procedural in nature, specifying sequences of steps the system should perform. Much of the knowledge in the Task Model is also specific to a particular mode of interaction (specifying, for example, the format of daily or weekly counseling conversations), while knowledge in the other models can be applicable to a wider range of computerized behavioral interventions, such

as tailored print or web pages [37] in addition to dialogue. However, Task Models can be specified in ways that promote reuse of their components, through appropriate parameterization and hierarchical decomposition of tasks and sub-tasks.

In summary, explicit representation of concept and task knowledge in standard formalisms facilitates reuse by identifying and exposing all the ways a computerized intervention can be parameterized (e.g., tailoring parameters, behaviors, etc.), what its inputs and outputs are (to facilitate interoperability), and how its components (task models and the data they depend on) can be extracted and recombined into new systems. This information can also be used to validate that an intervention has fidelity to a particular theoretical framework (as in [32]), or to provide a clinical summary of the theories, constructs and counseling techniques used in a given automated counseling session with a patient.

1.3. Overview

We have undertaken the development of a health behavior change counseling dialogue system that models the theory-driven therapeutic planning processes of a human health advisor during a counseling session, and is designed to be reusable. To address issues of reuse, we have built the system on public standard formalisms for computerized ontologies (OWL [38] and RDF [39]) and hierarchical task models (ANSI/CEA-2018 [40]), and have used these formalisms to explicitly represent some aspects of the knowledge described above. In the rest of this paper we review related work, then describe the behavioral medicine concept ontology and task modeling language formalism we are using. We focus our discussion on the design of the Task Model and how its hierarchical design and parameterization promote reuse. We then describe two health counseling dialogue systems created using these representations to demonstrate reuse. We finally present results from formative evaluation studies before concluding.

2. Related Work

Bickmore and Giorgino provide a review of work in health education and behavior change dialogue systems developed to date [18]. Most systems have been deployed as Interactive Voice Response systems over the phone, although a few have been developed using animated Embodied Conversational Agent interfaces. The majority of fielded systems have been designed for system-initiated conversation using state machines or hierarchical transition networks, implemented with ad hoc representations of persistent knowledge and dialogue structure. Application domains span exercise, diet, and smoking cessation promotion, medication adherence, and chronic disease self-care management education and promotion (e.g., for coronary heart disease and diabetes mellitus). The majority of the approximately two dozen systems evaluated in clinical trials have been shown to be effective.

Beveridge and Fox describe a speech-based dialogue system for clinicians, that performs breast cancer screening using an ontology of concepts in breast cancer together with a patient's medical record to drive the behavior of the system. However, this behavior is very narrow in scope (compared to the range of knowledge in Figure 1, for example), used only by the natural

language processing algorithms to allow clinicians to volunteer information or to state information using words and phrases not explicitly in the speech grammar.

Lenert, et al, describe an ontology of concepts from behavioral medicine, and use it, together with a task model in GLIF3 [41], to validate aspects of an exercise and diet intervention for adolescents [32]. Lenert's ontology is one of the first developed for behavioral medicine, describing health behavior change theories, constructs, protocols and their inter-relationships. However, his task model does not contain nearly enough detail to drive an automated intervention, only representing protocols to the "module" level, each potentially abstracting many turns of counseling dialogue (for example, it does not specify particular counseling techniques, or how the segments of a given counseling session should be sequenced). In addition, GLIF3 is both inappropriate and lacks sufficient representational power for dialogue planning. The GLIF languages were developed to encode clinical guidelines for physicians to follow, in which there is a single best sequence of actions for clinicians take in a given situation. This is fundamentally different from dialogue planning in which goals (system, patient, and mutual) must be explicitly represented and reasoned about, and task fragments are dynamically selected and added to the current task tree at each turn in the conversation.

3. Ontology of Behavior Change Counseling Concepts

Our ontology is a "domain ontology" [42, 43], which specifies the conceptualization of the domain of theory-driven, conversational agent-based behavior change interventions, capturing the kinds of knowledge outlined in Figure 1. Domain ontologies contain concrete categories that are of direct relevance to inference. A "core ontology" covering all of behavioral medicine has yet to be developed. Our ontology is "theory-driven" because only psychological states and therapeutic actions (such as counseling dialogue) that are sanctioned by a theory or method of health behavior change counseling are included in the model. We assume that the "counselor" (conversational agent) may have multiple conversations with a given "patient" (user) over time, and that its only means of effecting change is through dialogue. The ontology does not make any assumptions about the underlying dialogue model, and only makes reference to "Dialogue Actions" that can be realized as discourse segments consisting of one or more consecutive utterances by the counselor and patient. Thus, the knowledge in the ontology is not sufficient to produce a counseling dialogue; it must be augmented with information about how to produce specific utterances and how multiple utterances and Dialogue Actions should be sequenced. In our system, this information is encoded in our task model (Section 4).

We developed the ontology based on a review of concepts in behavioral medicine (e.g., [44]), prior work in computerized interventions by others [18], requirements for counseling dialogue systems we have developed (e.g., [45, 46]), review of prior behavioral medicine ontologies (notably [32]), and discussions with experts in behavioral medicine and computerized interventions. We also reviewed a number of "upper level" ontologies (e.g., [47]) for suitability, but decided that these representations were largely inappropriate for representing the psychological and linguistic concepts involved. We do build upon one existing ontology: the

OWL-time ontology [48]. Our ontology is developed in the Web Ontology Language (OWL) [38] using Protégé [49].

3.1. Core Concepts

Figure 2 shows the primary concepts in the ontology. These do not make any commitments to a particular health behavior change theory or technique, nor assume any particular behavior change target.

A **CounselingFramework** is essentially a collection of patient mental states and counselor therapeutic actions that can modify those states. This is more general and less restrictive than Lenert et al's behavioral medicine ontology [32], but allows us to index therapeutic dialogue actions based on what is known about the patient, and allows counseling methods, such as Motivational Interviewing [22], to be represented in addition to full-blown theories. Also, a counselor may perform many specific dialogue actions and consider many patient mental states in support of a theory, without the actions or the states being explicitly named constructs of the theory (e.g., the specific set of dialogue acts used assess a patient's stage of change, or used to enact a particular process of change for a given behavior). Actual behavioral medicine theories can be represented as more restrictive and fully defined subclasses of CounselingFramework, as shown in Figure 3 for the Transtheoretical Model [20].

An **Action** (in Figure 2) is something that an **Agent** (the automated counselor or the **Patient**) can perform, and a **TherapeuticAction** is something an Agent can do that is sanctioned by a CounselingFramework and appropriate in the context of a particular set of patient mental states to work towards a desired **InterventionGoal** (e.g., to increase physical activity). TherapeuticActions are either **TherapeuticDialogueActions**, performed during a counseling session, or **HomeworkActions**, performed by the Patient between counseling sessions. **TherapeuticDialogueActions** can be used to assess patient state, or to affect change in state. **HomeworkActions** are further decomposed into **HealthBehavior**s, involving the patient performing the health behavior of interest (e.g., "walking"), and **HomeworkActivity**, involving the patient doing a different activity in support of the InterventionGoal (e.g., signing up for a gym membership). **Tips** are a kind of ThereapeuticDialogueAction that are intended to help the patient resolve a particular **Obstacle** that is preventing him or her from reaching the InterventionGoal.

A counseling session will inevitably contain dialogue actions that are not sanctioned by any CounselingFramework, but are essential and expected parts of any human conversation. These **NonTherapeuticDialogueActions** include specifications of the overall structure of a conversation (e.g., greeting, followed by conversation body, followed by farewell), performance of ritualized behavior (e.g., specific greeting sequences, small talk), and interactional dialogue actions, such as acknowledgments.

Figure 2. Core Concepts in Ontology (Theory- and Behavior- Neutral)
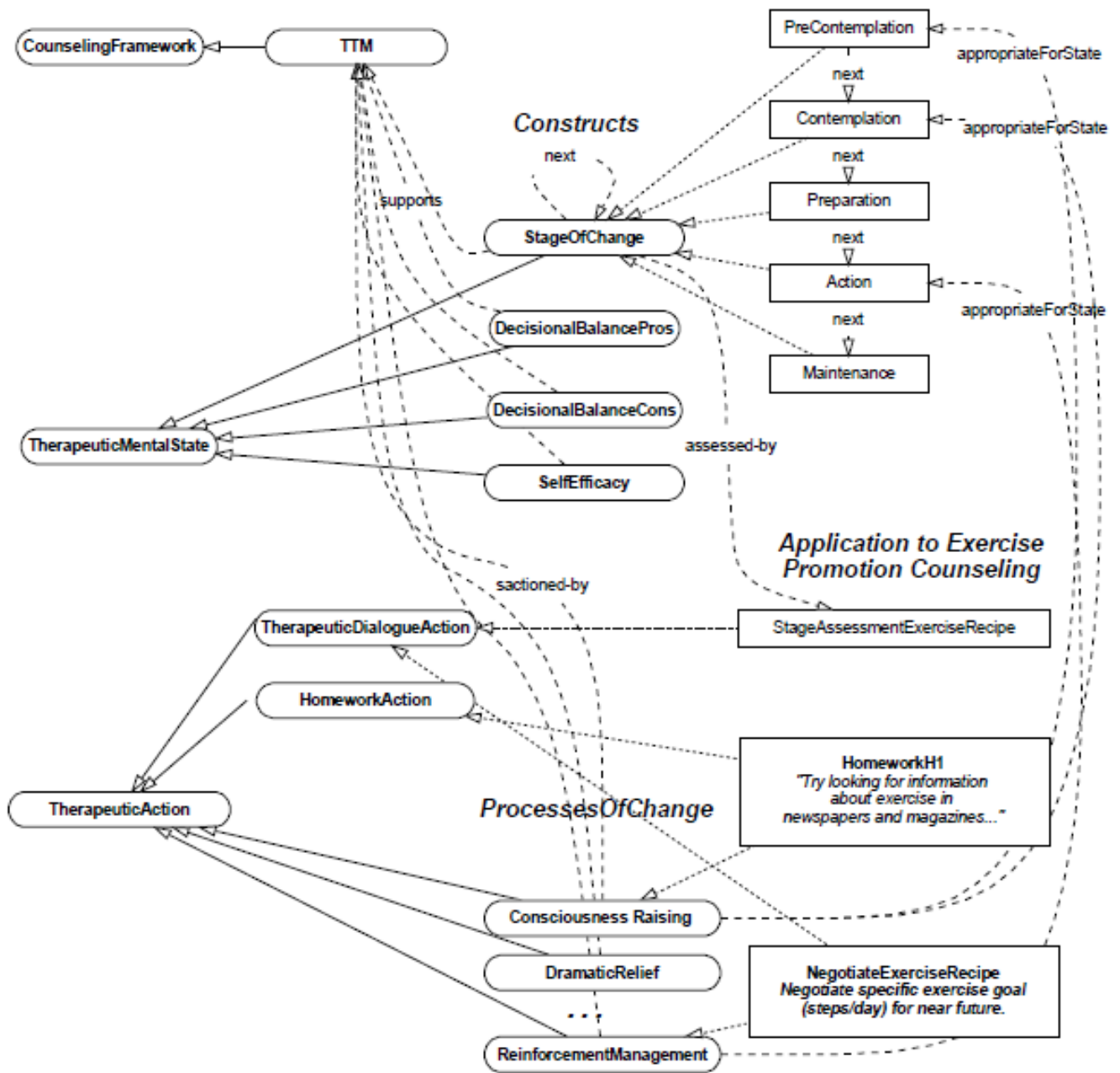**Ovals represent concepts; solid arcs represent subclass relations.**

Figure 3. Ontology Extensions for Transtheoretical Model Applied to Exercise Promotion Ovals represent concepts; boxes are instances. Solid arcs represent subclass relations; dotted arcs are instance relations.

Finally, a record of what happened during an **Intervention** consists of a sequence of **CounselingSessions**, each of which consists of a sequence of **ActionOccurrences** which record specific instances of DialogueActions (**DialogueActionOccurrence**) or HomeworkActions (**HomeworkActionOccurence**). This information allows the counselor to reflect on what has been done to date when planning TherapeuticActions, as well as to provide the basis for human-readable progress reports.

3.2. Example Extensions for Transtheoretical Model Applied to Exercise

Figure 3 shows extensions of the ontology presented in Figure 2 for the Transtheoretical Model (TTM, also known as the "stages of change model") [20, 61], and a few specific instances applying this extended ontology to exercise promotion counseling. TTM posits that individuals go through a set of distinct stages as they change, from Precontemplation (no intention to change), to Contemplation (intending to change but not acting yet), to Maintenance (well established habit), and that individuals need different interventions ("processes of change") based on the stage they are in. The constructs of the CounselingFramework **TTM** include **StageOfChange**[1] (with the individual stages as instances), **DecisionalBalance** (with **Pros** and **Cons** as instances), and **SelfEfficacy**. TherapeuticActions are decomposed into the ten broad TTM categories of "processes of change", including **ConsciousnessRaising** (discovering new information that supports the desired health behavior change), **DramaticRelief** (experiencing negative emotions that accompany unhealthy behavior), and **ReinforcementManagement** (rewards for positive health behavior change).

Figure 3 also includes a few examples of specific TherapeuticAction instances for TTM-based exercise promotion. **StageAssessmentExerciseRecipe** indexes a "recipe" in the Task Model (Section 4.1) that asks the patient a series of questions to assess their StageOfChange. **HomeworkH1** is an example of a homework assignment in which the patient is asked to investigate information about the benefits of exercise in various media (an instance of ConsciousnessRaising). In our system, this literal is used to index a series of recipes in the Task Model that propose, negotiate, and follow up on the homework assignment, depending on the discourse context. Finally, **NegotiateExerciseRecipe** indexes a recipe that negotiates a specific exercise goal (in steps per day of walking) with the patient.

4. Task Model for Health Behavior Change Counseling

For a health counseling dialogue system, the Task Model describes the structure of a counseling session by describing each of the atomic steps in the dialogue (system or patient utterances) and the relationships among them. To model counseling dialogue in an extensible and reusable manner, we use a standard declarative task representation language (CEA-2018 [40, 50]) for

---

[1] We note that TTM "stage of change" is actually comprised of both physical (time performing criterion behavior) and mental (intention to perform criterion behavior) states, but for the purpose of planning counseling dialogue it can be treated as a strictly mental construct.

representing dialogue fragments, and have developed our own dialogue planner (DTask [51]) that uses these fragments to enact a dialogue with a patient.

This design was motivated by our previous experience in building systems that required authoring large quantities of dialogue [52, 53]. In these systems, we modeled dialogue using hierarchical transition networks, typically with agent utterances as states in a network and patient utterances as transitions. This formalism had the advantage of being simple to understand, particularly for clinicians and behavioral scientists without backgrounds in computer science or computational linguistics. However, we observed over the course of building several systems designed for health counseling, that reuse and generalizability of the dialogue between applications was often poor. In addition, while the therapeutic dialogue of these systems was based on well-established counseling theory and practice, this underlying knowledge was not explicitly modeled in the structure of the dialogues. Our current approach has made it possible to create more general and reusable dialogue models (e.g., through task and ontology abstractions). At the same time, parts of the task library can be used to codify more domain knowledge into the declarative dialogue model representation.

4.1. Task Model Representation Language: CEA-2018

Following Shared Plans theory [54, 55], we treat dialogue as a collaboration in which participants coordinate their action towards achieving a shared goal. The intentional structure of dialogue is modeled as hierarchical task decomposition: a library of recipes (goal decompositions), which may be used to achieve the overall goal of the collaboration.

Our dialogue recipes are specified in a declarative language that was recently approved as an ANSI Consumer Electronics Association standard: CEA-2018 [40, 50]. CEA-2018 provides a number of abstractions that are useful in designing reusable task models including:

- Task ("applicability") preconditions, enabling many recipes to be developed for accomplishing a given goal, with conflict resolution performed at runtime.
- Partial ordering of recipe steps, with complete ordering as default.
- Optional recipe steps.
- Declaration of input and output parameters for a given recipe, providing a locally scoped parameter passing mechanism.

CEA-2018 includes ECMAScript [56] for performing algorithmic tasks, data representation, and providing logical expressions for applicability preconditions.

To more easily describe dialogue actions, we extended the CEA-2018 specification with a declarative specification of dialogue based on the notion of an "adjacency pair" which is a pair of adjacent utterances that are logically related (e.g., a question followed by a response) [57]. It consists of an agent utterance and a multiple choice list of client responses; this template comprises a primitive action (i.e. the lowest levels of the tree) in the goal description. An adjacency pair template achieves a particular sub-task in a recipe; there may be an arbitrary number of adjacency pair templates, as well as recipes, which can achieve a goal. Furthermore,

each template automatically includes a patient response for clarification of the agent utterance (e.g., "Could you repeat that please?"). Clarifications either reproduce the same agent utterance, or an alternative utterance, if available.

A more refined set of classes of adjacency pairs (such as questions/answers, assertions/acknowledgements, etc) from which individual adjacency pairs would be represented has not been created because there is no semantics of natural language utterances powerful enough to capture these distinctions in a general way. Without such a semantics, individual adjacency pairs can only be specified in the manner developed in this work.

An adjacency pair template specifies agent and patient natural language utterances, but each utterance may include variables that are to be filled in when the utterance is produced ("template-based text generation" [58]), such as the "<desired>" variable in Figure 7. Utterance text can also be annotated with XML tags to specify, for example, prosodic or nonverbal behavior to be used in its performance by the system [59].

4.2. Task Model Interpreter: DTask

DTask is a hierarchical task decomposition-based dialogue planner, inspired by the COLLAGEN collaborative dialogue system [60]. It is designed to model and execute system-directed dialogue, where patient input is performed via selection from a multiple choice menu of possible utterances, dynamically updated at each turn of the dialogue. DTask uses a CEA-2018 recipe library to plan a task decomposition for each segment of a dialogue.

DTask provides support for loading and saving data and metadata represented in Resource Description Framework (RDF) [39] at runtime. RDF is a general-purpose knowledge representation format, distinguished by the use of a Universal Resource Identifier (URI) to identify each resource or individual being described, independent of a particular RDF document or knowledge store. Several ontology languages, such as the Web Ontology Language (OWL) are available for describing the structure of a RDF knowledge store, and for enabling inference upon them. A large ecosystem of software, both open-source and closed-source, exists for working with RDF, including editors, storage engines, and query and inference engines.

Arbitrary RDF documents can be loaded into the DTask knowledge store, either from files or a URL using an ECMAscript-based API. This function is currently used to load parts of the ontology into the runtime dialogue system and to save updated parts of the ontology (User Model) at the completion of a counseling session.

DTask can also generate an RDF description of a loaded set of recipes, including a unique identifier (URI) for each goal, goal decomposition, and dialogue turn, and a description of the goal-subgoal relationships. These descriptions can be annotated with arbitrary additional metadata, such as the behavior change theory a given recipe derives from. DTask can optionally use this metadata to filter a set of recipes, only including those that meet a specified set of criteria, for example only loading those models that adhere to a particular theoretical model. These models are useful in design and reuse by providing a visualization of the goals and recipes in a given dialogue system (Figure 4).

During the course of a dialogue, DTask maintains a representation of a "plan tree", recording the recipes that are actively being used and their goal-subgoal task decomposition relationships (as in [60]). This representation is also maintained in RDF and can be exported at the completion of a dialogue. An advantage of this approach is that the exported RDF provides a complete description of the counseling session. This description can include not only a complete transcript of the conversation (as in the Appendix), but the rationale for why the system used any particular utterance. Such a description makes it possible for a professional to assess not only how a given person is proceeding in their counseling but also how the computer system is accomplishing its counseling tasks. Just as many human counselors are now required to provide detailed notes about their sessions with clients, so this system can do likewise in a readable and useful form.
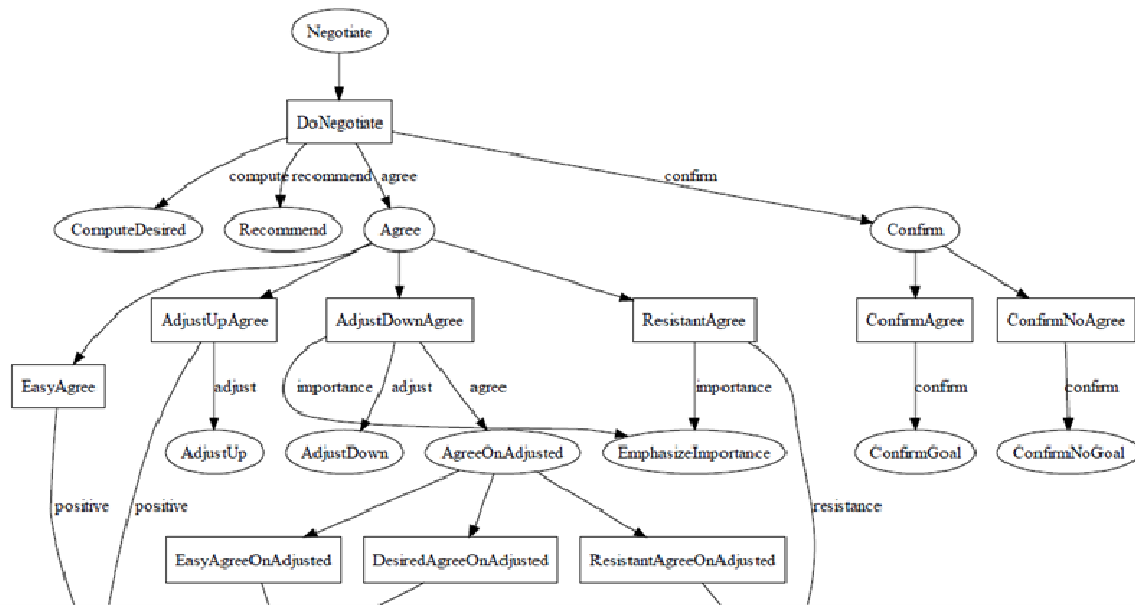

Figure 4. Fragment of the Task Model Hierarchy for Negotiating Goals

Boxes represent recipes; ovals represent goals.

In summary, the DTask system makes two improvements upon the framework of COLLAGEN [60]. The first of these is the use of adjacency pair templates, which are essential for providing the means of individual utterances and clarifications for users with agent follow up. COLLAGEN relied on a weak but general semantic form which did not provide the level of detail that is available with the adjacency pair template mechanism. The second improvement is the DTask Planner, which provides a mechanism for choosing what to do or say next, once the current state of the dialog is updated given user input. In the COLLAGEN framework, the choice of what to do after interpreting user input, is determined by set of rules (implemented as

plug-ins) about how to proceed given the current state of the dialog. The default planner using no plug-ins chooses the next action (spoken or otherwise) that the agent can do. DTask goes about planning a task to take advantage not only of the task hierarchy (as COLLAGEN does) but also to access metadata in the RDF framework.

4.3 Integration of Semantics in the Ontology and Task Models

The ontology described in Section 3 and Figures 2-3 is used to specify the relationships among constructs from behavioral medicine theories and methods and the recipes that can either assess the construct for a given user or work towards an intervention goal given the assessed values for a set of constructs. The ontology is also used to specify the semantics for certain classes of therapeutic utterances, such as patient utterances in motivational interviewing (described in Section 4.5.2), and values that parameterize the behavior of the automated counseling system, such as the patient's name and strings and parameters that describe the intervention target behavior (e.g., "walking steps" vs. "servings of fruit and vegetables"). Additional semantics are specified and used entirely within the task models if they are local in scope (used by a set of related recipes) and not saved persistently across counseling sessions, for example the patient response to an initial stage of change question ("Do you currently exercise at least five days per week, for thirty minutes or more?") that is used to determine which follow up question should be asked.

4.4. Health Counseling Task Model Design

Our current counseling Task Model is based on the Transtheoretical model [20, 61], social cognitive theory [21], and motivational interviewing [22]. We initially focused our effort on those counseling techniques relevant for exercise (walking) promotion, and developed models to address all stages of change, including Precontemplation, an area most automated interventions do not focus on.

Our approach to Precontemplation, Contemplation and Preparation is drawn from Motivational Interviewing, a brief, directive, client-centered counseling method for enhancing intrinsic motivation to change by helping clients explore and resolve ambivalence [22]. The method includes a number of conversational strategies for eliciting "change talk" from clients in order to increase their motivation to and confidence in changing their behavior. While Motivational Interviewing can help individuals progress through all stages of change, it is especially important to build motivation and confidence in the earliest stages.

Our approach to the later stages of change involves long-term and short-term goal setting negotiation, positive reinforcement when goals are met and "problem solving" to overcome barriers when goals are not met (social cognitive theory). We also provide stage-appropriate homework assignments, tips and educational information throughout every intervention.

In addition to drawing from the counseling literature, we videotaped several counseling sessions between a clinical psychologist specializing in health behavior change and his patients. Transcripts of these interactions were used as the basis for many aspects of our system, from the

overall conversational structure to specific recommendations and homework assignments. This psychologist also provided a significant amount of input to and feedback on our designs.

4.5. Task Model Implementation

Our goals in modeling the task structure of health behavior change counseling sessions were to maintain as much fidelity as possible to counseling practice while creating abstractions that promoted reusability. The system's knowledge of health behavior change counseling is represented in the ontologies outlined in Figures 1-3, loaded into the system from RDF files and represented in the runtime system as ECMAscript objects.

4.5.1. High-level Structure of a Counseling Dialogue

Figure 5 shows a fragment of the task models which represent the top-level steps in a typical brief counseling dialogue. These steps include (line numbers reference the sample dialogues in the appendix):

```
task Conversation
    input parameters: behavior
    precondition: (applicable when session > 1)
    steps:
       1. Opening
       2. RitualSocial
       3. ReviewTasks
       4. Assess
       5. Counseling
       6. AssignTasks
       7. PreClosing
       8. Closing
    ...
```

Figure 5. Task model pseudocode showing the high-level structure of a counseling dialogue.

1. opening the conversation with a greeting (2-1 to 2-2, and 3-1 to 3-2),
2. conducting ritualized social dialogue, including small talk (e.g., about the weather) and inquiries into the patient's general emotional and physical state (e.g., "how are you?") (2-3 to 2-12, and 3-3 to 3-16),
3. reviewing previously assigned behavioral tasks and goals (3-17 to 3-26),
4. assessing the patient's current state in the intervention, including stage of change and attitude toward the behavior (2-13 to 2-18, and 3-27 to 3-32),
5. conducting counseling based on Motivational Interviewing (2-19 to 2-37, and 3-33 to 3-37),
6. negotiating new behavioral goals and tasks (2-38 to 2-54),

7. preparing to close the conversation [62] (2-55 to 2-58), and
8. delivering a farewell (2-59 to 2-60).

We note two particular features of this recipe. First, while some portions (e.g., the opening, social dialogue, pre-closing and closing) are applicable to dialogue in a wide variety of domains, much of the recipe is specific to counseling dialogue. Second, while we do not claim that this recipe represents the only possible decomposition of a counseling dialogue, we argue that this recipe produces a more highly coherent counseling dialogue than many other possible recipes, due to the particular ordering of subtasks. For example, reviewing previous tasks should be done near the beginning of a dialogue, as this represents a continuation of a previous conversation. Both reviewing tasks and assessing the patient's current attitudes should be done before assigning new tasks, as both steps provide information that can be used to more effectively select new counseling tasks.

4.5.2. Structure of Motivational Interviewing Dialogue

In Motivational Interviewing, a counselor typically uses open-ended questions to elicit statements from the client relevant to their motivation or confidence to engage in the target behavior. The counselor can then respond with tactics designed to elicit further statements, such as requesting more information (e.g., "Tell me more…") or reflection (i.e., restating the client's utterance). Lines 2-21 to 2-37 of the appendix give an example of this type of dialogue. In our system these tactics are modeled in a reusable manner, so that new target behaviors, and new topics of discussion relevant to those behaviors, can be easily added. We follow the abstraction strategies discussed above, and create abstracted high-level recipes that do not contain information about any particular behavior or topic. A key challenge in this approach is defining an appropriate semantics for statements made by the patient during counseling.

Miller and Rollnick [22] discuss the tactics that a counselor can use to respond to two different types of patient statements: "change talk" indicates motivation for behavior change and/or confidence to change, while "resistance" indicates a lack of motivation and/or confidence. They also present 4 categories of meaningful statements, all of which can be positively valenced as change talk, or negatively valenced as resistance:
1) Positive or negative implications of the status quo.
2) Positive or negative implications of the proposed behavior change.
3) Statements of positive or negative outlook towards change.
4) Statements of intention or lack of intention to change.

The first two categories generally indicate motivation, or lack of motivation for change, while the latter two indicate confidence, or lack of confidence.

We have adapted this taxonomy to provide a simple domain-specific semantics for patient statements during the Motivational Interviewing portions of the dialogue. A patient statement is characterized by a tuple of three values:

1) *valence*: change talk or resistance
2) *category*: one of the 4 categories discussed above
3) *content*: the specific topic being discussed

These are represented in the ontology as properties of MITherapeuticDialogue action, a subclass of TherapeuticDialogueAction.

Figure 6 gives an example of the use of these semantics in an adjacency pair template. This example shows an adjacency pair which can be used if the agent chooses to respond to a patient utterance about stress with a reflection tactic. An example dialogue segment would be the following:

1. Patient: I'm stressed all the time.
2. Counselor: You sometimes feel like you have too much stress.
3. Patient: If I did exercise, it might be better.

The patient utterance at (1) would be categorized with a valence of "change_talk", a category of "status_quo_implication", and content of "stress", triggering use of the recipe in Figure 6 and the generation of utterance (2), and an offer to the patient to respond with utterance (3). There are currently over 400 such adjacency pairs specified for the "Reflection" tasks and for similar tasks representing other possible tactics a counselor may use. While this quantity of dialogue represents a significant authoring effort, such adjacency pairs are straightforward to write and can be extended by domain experts with little computer science knowledge or knowledge of the overall dialogue model. Conversely, there is a much smaller set of higher-level recipes which determine the appropriate tactic to use in order to respond to a particular patient utterance. These high-level recipes are more complex and do require computer science expertise to create. However, the high-level recipes are independent of the particular topics discussed, and therefore these two portions of the dialogue model can be modified and extended independently of each other.

To summarize, the treatment of Motivational Interviewing dialogue given here illustrates how the careful specification of a domain-specific semantics allows the separation and abstraction of high-level recipes from low-level dialogue. This abstraction enables the creation of complex dialogue models which can be extended on multiple levels: with new target behaviors and topics of discussion, and with new higher-level tactics for counseling.

```
task Reflection
  input parameters: behavior, talk_valence,
                    talk_category, talk_content
  outputs: valence, category, content
  adjacency pair:
    precondition: applicable when
       behavior='exercise' and
       talk_valence='change_talk' and
       talk_category='status_quo_implication' and
       talk_content='stress'
    agent utterance:
       You sometimes feel like you have too much
       stress.
    user utterance:
       If I did exercise, it might be better.
       valence ← 'change_talk'
       category ← 'change_implication'
       content ← 'stress'
…
```

Figure 6. Example recipe for motivational interviewing

4.5.3. Structure of Social Cognitive Counseling Dialogue

Once a patient is past the Contemplation stage, the agent begins negotiating long-term and short-term explicit behavioral goals (e.g., number of steps walked). The overall objective is to gradually move a patient from their current level of performance up to the criterion level. Goals are negotiated by giving patients some latitude in selecting their short-term (day-to-day) goals to account for fluxuations in their motivation, self-efficacy or personal situation (e.g., going away for a weekend).

Once a goal is established, patient performance relative to their goal is reviewed in subsequent counseling sessions. If they have met their goal, the agent provides positive reinforcement ("great job!"). We have observed (as have other researchers [63]) that patients do acquire a sense of accountability to an automated counseling agent that motivates them to meet their goals. If goals are not met, the agent engages the patient in a problem-solving discussion to identify a specific barrier or obstacle to meeting their goals and makes suggestions for how the patient might overcome it.

4.6. Task Model Design for Reuse

The use of a hierarchical task decomposition as a model of dialogue permits the reuse of dialogue structures at different levels of abstraction. For example, we observed that several points in the dialogue required the agent to negotiate a numerical quantity with the patient, such as a long-term goal (e.g., an average number of steps to walk after several weeks) or a daily goal (e.g., a number of steps to walk before the next session). Although there may be many

differences between the two dialogue fragments, the high-level steps are the same: the agent makes an initial recommendation, and depending on the patient's answer, may respond with a higher or lower daily goal, or (in the case of strong resistance), a discussion of possible obstacles. Figure 7 shows a portion of a high-level recipe for negotiation, and an example of a dialogue turn that works with it. To extend the system to perform negotiation on long-term diet goals (for example), requires only adding a new adjacency pair for the "Recommend" task (and other similar tasks) which is applicable only for that behavior.

```
task Negotiate
    input parameters: behavior, target
    outputs: goal
    steps: 1. ComputeDesired
           2. Recommend
           3. Followup
           4. Confirm
task Recommend
    input parameters: behavior, target, desired
    outputs: response
    precondition: (behavior='exercise' and
                        target ='long_term_goal'
    adjacency pair:
    agent utterance:
        Based on your activity level, I suggest
        gradually working up to ‹desired› steps per
        day over the next eight weeks
…
```

Figure 7. Example pseudocode for a high-level recipe and a low-level dialogue specification

5. Evaluation

We have developed two behavioral interventions using the above framework: the first to promote physical activity (walking) and a second to promote fruit and vegetable consumption. In this section we describe separate, formative evaluation studies of each system. In addition, we report on the amount of reuse between two behavioral interventions. Both evaluations were approved by the Northeastern University IRB, and informed consent was obtained at the beginning of each session by a research assistant.

5.1 Evaluation of Exercise Promotion Intervention

We conducted a mixed-method, formative evaluation of the dialogue system described above in the domain of exercise promotion. Study participants interacted with the system via an animated

Embodied Conversational Agent interface (Figure 8), making their contributions to the dialogue via a touch screen monitor [51]. Three evaluation methods were used in parallel: 1) evaluation of dialogue transcripts by a health behavior change expert; 2) self-report assessments of each dialogue by participants following completion of an entire counseling session; and 3) self-report assessments of every dialogue utterance by retrospective review (playback) of each dialogue by participants.
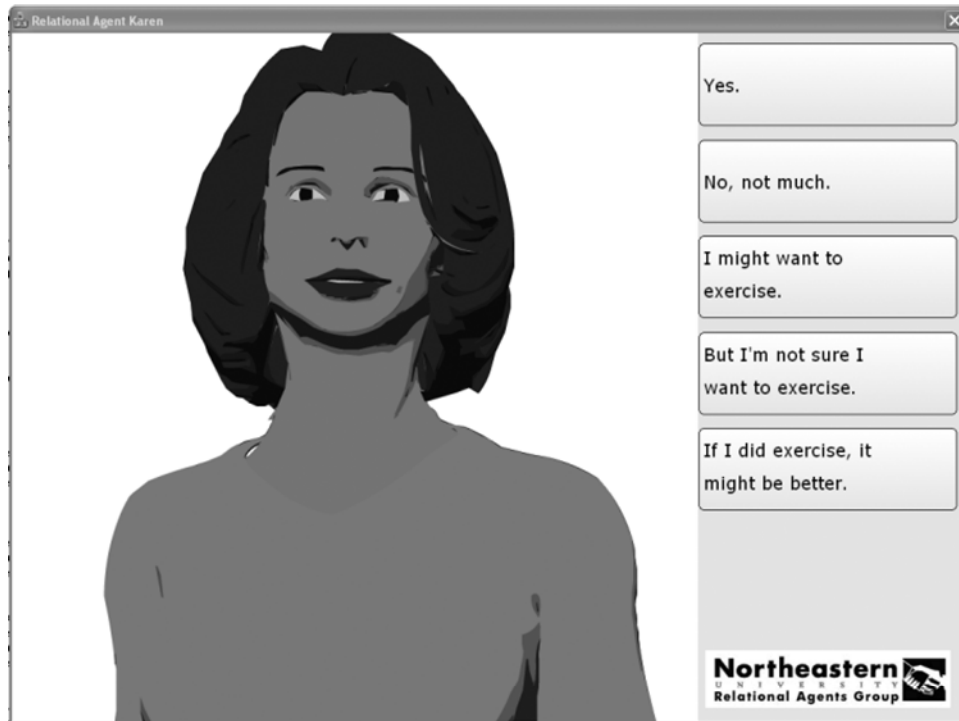

Figure 8. Embodied Conversational Agent Interface

5.1.1 Participants and Protocol

Eight participants, aged 21-41 (mean 30.4), 75% female, were recruited via an online advertisement: three were in preparation, three in action and two in maintenance relative to the stages of change for minimum recommended levels of physical activity [20].

Each participant conducted an initial conversation with the agent then filled out a summative evaluation questionnaire. They were then told to pretend that they had come back the next day to conduct a follow-up conversation with the counselor. Immediately following this second conversation, the participant filled out another summative evaluation questionnaire then conducted a retrospective review of the entire second conversation. In this review, they watched a computer-controlled replay of their conversation that stopped after each turn and asked a series of questions. They then repeated this protocol for a third and final conversation.

5.1.2 Measures

An expert reviewer trained in Motivational Interviewing, but who was not involved in the design of the intervention, provided summative evaluations of text transcripts of nine conversations (three each from the participants in Preparation stage). She used the global ratings of Empathy and Motivational Interviewing Spirit (each on 7 point scales) from a standard evaluation instrument used to assess human counselor fidelity to Motivational Interviewing principles and practice [64].

The summative self-report measures by participants comprised 19 scale questions (range 1-7), assessing overall satisfaction (3 items, Cronbach's alpha 0.88), agent empathy (3 items, alpha 0.82), Motivational Interviewing "spirit" (8 items, alpha 0.86), and relationship with the animated counselor (3 items, alpha 0.83).

During the retrospective review, following every agent utterance, participants were asked to rate the appropriateness of the utterance, the naturalness of the utterance, and their trust in the agent, all on 7 point scales. Following every replay of a participant contribution to the dialogue, they were asked to rate how well their choice reflected what they actually wanted to say on a 7 point expressivity scale.

5.1.3 Results

Table 1 presents descriptive statistics of the study results, with retrospective results averaged over entire conversations. Expert ratings of agent empathy and Motivational Interviewing spirit were surprisingly high, and the expert provided insightful justification for her scores. The first conversation is primarily structured to provide patients with an orientation to the agent and intervention and to perform a stage-of-change assessment. The expert felt that this conversation did not elicit much information from the study participants (also reducing empathic opportunities), resulting in its relatively low score. By the third conversation, however, the expert felt that the conversation was being tailored to both the participant's current needs as well as to things talked about in the prior sessions, resulting in higher scores.

Overall ratings of agent appropriateness and participant expressivity in the retrospective review were very high, with less than 3% of agent utterances and less than 4% of participant utterances rated below the midpoint of the scales. However, this methodology did enable us to identify parts of the dialogue that had an especially negative impact on participants. For example in one sequence, the agent gave appropriate empathic feedback to a participant for not feeling well, then immediately launched into social chat by saying "So, I hope your day is going well.", which did not impress the participant (turn 7 in Figure 9).

Table 1. Exercise Intervention Evaluation Ratings (mean (SD))

| Expert Summative (N=3 participants)	Dialog 1	Dialog 2	Dialog 3
Agent Empathy	5.00 (0.00)	5.33 (1.15)	6.00 (1.00)
MI Spirit	4.33 (0.58)	5.00 (1.00)	6.00 (1.00)
User Summative (N=8 participants)			
Satisfaction	5.33 (1.38)	5.57 (1.07)	4.92 (1.43)
Agent Empathy	5.04 (1.34)	5.38 (1.03)	4.92 (1.43)
MI Spirit	4.84 (1.14)	5.43 (0.99)	4.55 (1.37)
Relationship	4.71 (1.29)	5.24 (0.76)	4.88 (1.18)
User Retrospective (N=8 participants)			
Appropriateness		6.40 (0.54)	6.35 (0.49)
Naturalness	*Not*	6.35 (0.67)	6.00 (0.73)
Trust	*measured*	6.13 (0.91)	6.01 (0.76)
Expressivity		6.23 (0.66)	6.02 (0.79)

All 7-point Scales with 7.0 Best
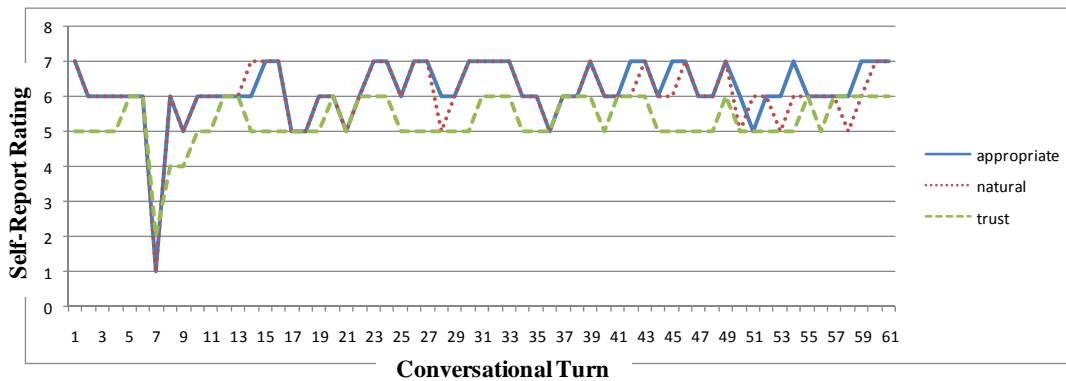
Figure 9. Example Turn-by-Turn Retrospective Review

5.2 Evaluation of Reuse

We extended the DTask task models to promote fruit and vegetable consumption to demonstrate that the abstractions used reliably lend themselves to reuse. These extensions required an additional 4 tasks (98% reuse), 3 recipes (98% reuse), and 961 agent utterances (14% reuse), totaling an additional 24,096 lines of source code (22% reuse). This demonstrates that the high level task abstractions (recipes and goals encoding general counseling strategies) are highly reusable, and most of the effort in porting the system to a new health behavior involves authoring new agent utterances at the leaves of the task hierarchy. Extending the system to diet promotion took 9% of the calendar time and 4% of the person hours required to develop the initial exercise promotion system.

5.3 Evaluation of Fruit and Vegetable Promotion Intervention

We conducted a mixed-method, formative evaluation of the fruit and vegetable promotion dialogue system, using the same protocol and measures used in the exercise promotion system evaluation described in section 5.1 above.

5.3.1 Participants

Nine participants, aged 22-68 (mean 41), 78% male, were recruited via an online advertisement. With respect to the behavioral criteria of 5 or more half-cup servings of fruit and vegetables on five or more days of the week [17], two were in pre-contemplation, two were in contemplation, three were in preparation, one was in action and two were in maintenance.

5.3.2 Results

Table 2 presents descriptive statistics of the User Summative and Retrospective and Expert Summative results. These results are very close to those for the exercise promotion system, indicating that participants were as satisfied with the fruit and vegetable promotion system and had as positive attitudes towards the automated counselor as they did with the exercise promotion system.

6. Limitations

There are many limitations of our overall approach, particular implementation, and evaluation strategy. Regarding our approach, there is admittedly somewhat of a disconnect between the ontology described in Section 3 and the task models, described in Section 4. Our approach is centered on the task models, with certain key semantics specified in the OWL ontology and represented in RDF. Ultimately, these two representations should be merged, with the task models themselves, and all semantic representations, represented in the ontology. One particular shortcoming is that conversational goals, which drive the behavior of the dialogue system, are currently not represented at all in the ontology.

Table 2. Fruit and Vegetable Intervention Evaluation Ratings (mean (SD))

Expert Summative (N=3 participants)	Dialog 1	Dialog 2	Dialog 3
Agent Empathy	5.33 (0.58)	6.00 (0.00)	5.50 (0.71)
MI Spirit	4.00 (0.00)	5.00 (1.00)	6.00 (1.41)
User Summative (N=9 participants)			
Satisfaction		5.15 (1.19)	5.04 (1.53)
Agent Empathy	*Not*	4.22 (1.49)	4.79 (1.83)
MI Spirit	*measured*	4.65 (0.76)	4.73 (1.00)
Relationship		4.48 (0.80)	4.58 (1.32)
User Retrospective (N=9 participants)			
Appropriateness		6.02 (0.55)	6.17 (0.76)
Naturalness	*Not*	6.02 (0.62)	5.97 (0.88)
Trust	*measured*	5.64 (0.91)	5.62 (0.88)
Expressivity		5.90 (0.59)	6.27 (0.66)

Our use of template-based text generation, while conceptually simple and straightforward to implement, limits the scalability and reuse of the system, since many utterance templates must be manually re-written for any application that is not very similar to the application on which it is based. This is exemplified by our only realizing 14% reuse of utterances when porting the exercise promotion system to fruit and vegetable promotion, compared to 98% reuse of the other knowledge. A text generation module that generates natural language from first principles (lexicon, grammar, etc. [58]) could significantly improve the re-usability of our approach. For example, all utterances could be automatically tailored based on patient demographics and the degree of familiarity between the patient and the counselor; for now such assumptions must be hard-coded into the utterance templates.

Our formative evaluation studies also have several limitations. First, the sample sizes were very small, limiting the generalizability of the results, although they are typical for usability studies that are intended primarily to guide the design of a system [65]. Second, the studies lacked some aspects of ecological validity, since participants were asked to pretend that a day had elapsed between each session, when in fact it had only been a few minutes. This compressed time may have led to an increased sense of repetitiveness, since participants would be more likely to notice repeated elements of dialogue, leading to lower satisfaction ratings. Also, providing hypothetical feedback (e.g., about whether they had actually gone for a walk or not)

may have led to somewhat different ratings than would have been provided in a real intervention situation. Finally, the evaluations did not provide true validation of the approach in terms of long-term health behavior change in comparison to a control condition. The exercise and fruit and vegetable interventions we have developed will be evaluated in a randomized 4-arm, 60-day trial to compare each intervention to a combined intervention in which both exercise and diet modifications are promoted, and with a non-intervention control group.

7. Conclusion and Future Directions

We have demonstrated that health counseling dialogue can be modeled using representations that promote reuse, that study participants and expert health counselors are largely satisfied with the results and their fidelity to health behavior change theory, and that the representations do indeed lend themselves to reuse.

The two behavioral interventions we implemented were intentionally selected to be similar at the most abstract levels; both are acquisition interventions, and both are behaviors that can be incrementally modified once a patient's motivation to change has been established. Extending our models to a very different health behavior change problem, such as smoking cessation, would require much more extensive modifications to the models. Even for those problems, though, much could be reused; the overall structure of a counseling session, the structure of motivational interviewing dialogue, and many smaller dialogue fragments (such as greetings and social chat) would likely remain unchanged. In addition, other kinds of modifications, such as using Motivational Interviewing with patients in all stages of change (not just those in precontemplation or contemplation) would be very easy, requiring only that the applicability conditions for the Motivational Interviewing recipes in question be changed (assuming the dialogue content itself could be re-used across stages without modification).

The DTask dialogue engine, the LiteBody web-based Embodied Conversational Agent, and the ontology and task models for the exercise and fruit and vegetable behavioral interventions will all be released open source in 2010 [66]. We hope that these resources will, at a minimum, promote further discussion and development and reusable tools for building computerized behavioral interventions. The ultimate evaluation of this work will be the extent to which these tools are reused by other researchers and the extent to which they facilitate dissemination. Through the contributions of many researchers in an open source framework, we could eventually deploy systems that can be tailored on hundreds of variables rather than the handful most systems use today [37].

The continued development of ontologies for behavioral medicine is an important activity in its own right, as it can help to drive the many disparate theoretical camps to consensus on what works, or at least provide a framework in which theories can be more precisely compared and evaluated. As Lenert noted, the development of such formalisms helps achieve many goals, such as making the linkages between protocols and theories more explicit, and aiding in the evaluation of protocols and protocol components [32]. By extending these formalisms to the point at which they can actually drive the behavior of a computerized intervention, the subjectivity in mapping theories from behavioral medicine to interventions can be greatly reduced, improving our ability

to compare and evaluate theories. We also hope that this work serves to promote collaboration within and across the fields of medical informatics, behavioral medicine, computational linguistics, and AI.

Acknowledgements

References

[1] Mokdad AH, Marks JS, Stroup D, Gerberding J. Actual causes of death in the United States, 2000. JAMA 2004;291: 1238-45.

[2] U.S. Dept of Health and Human Services. Physical activity and health: a report of the Surgeon General. In. Atlanta, GA: U.S. Department of Health and Human Servies,Centers for Disease Control and Prevention, National Center for Chronic Disease Prevention and Health Promotion.; 1996.

[3] Bouchard C, Stephens T. Physical activity and health: international proceedings and consensus statement. Champaign, IL: Human Kinetics; 1994.

[4] Stewart A, Haskell W. Endurance exercise and health-related quality of life in 50-65 year old adults. The Gerontologist 1993;33: 782-789.

[5] LaPorte R, Adams L, Savage D, Brenes G, Dearwater S, Cook T. The spectrum of physical activity, cardiovascular disease and health: an epidemiologic perspective. American Journal of Epidemiology 1984;120: 507-571.

[6] Young D, Haskell W, Jatulis D, al e. Associations between changes in physical activity and risk factors for coronary heart disease in a community-based sample of men and wormen: The Stanford Five-City Project. 138 1993;4.

[7] Services USDoHaH. The Surgeon General's Report on Nutrition and Health. In. Washington, DC; 1998.

[8] McGinnis J, Foege W. Actual causes of death in the United States. JAMA 1993;270: 2207-2212.

[9] Knekt P, Reunanen A, Javinen R, et al. Antioxidant vitamin intake and coronary mortality in a longitudinal population study. Am J Epidemiol 1994;139: 1180–1189.

[10] Gillman M, Cupples L, Gagnon D, et al. Protective effect of fruits and vegetables on development of stroke in men. JAMA 1995;273: 1113–1117.

[11] Gaziano J, Manson J, Branch L, et al. A prospective study of consumption of carotenoids in fruits and vegetables and decreased cardiovascular mortality in the elderly. Ann Epidemiol 1995;5: 255–260.

[12] Key T, Thorogood M, Appleby P, et al. Dietary habits and mortality in 11,000 vegetarians and health conscious people: results of a 17 year follow up. BMJ 1996;313: 775–779.

[13] Rimm E, Ascherio A, Giovannucci E, et al. Vegetable, fruit, and cereal fiber intake and risk of coronary heart disease among men. JAMA 1996;275: 447–451.

[14]   Ness A, Powles J. Fruit and vegetables and cardiovascular disease: a review. Int J Epidemiol 1997;26: 1-13.
[15]   Joshipura K, Ascherio A, Manson J, et al. Fruit and vegetable intake in relation to risk of ischemic stroke. JAMA 1999;282: 1233–1239.
[16]   Li R, Serdula M, Bland S, et al. Trends in fruit and vegetable consumption among adults in 16 US states: Behavioral Risk Factor Surveillance System, 1990-1996. Am J Public Health 2000;90: 777-781.
[17]   Krauss R, Eckel R, Howard B, aAppel L, Daniels S, Deckelbaum R, et al. AHA Dietary Guidelines: revision 2000: A statement for healthcare professionals from the Nutrion Committee of the American Heart Association. Stroke 2000;31: 2751-6.
[18]   Bickmore T, Giorgino T. Health Dialog Systems for Patients and Consumers. J Biomedical Informatics 2006;39: 556-571
[19]   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: 21-8.
[20]   Prochaska J, Marcus B. The Transtheoretical Model: Applications to Exercise. In: Dishman R, editor. Advances in Exercise Adherence. Champaign, IL: Human Kinetics; 1994, p. 161-180.
[21]   Bandura A. Social foundations of thought and action: a social cognitive theory. Englewood Cliffs,NJ: Prentice-Hall; 1986.
[22]   Miller W, Rollnick S. Motivational Interviewing: Preparing People for Change. second ed. New York: Guilford Press; 2002.
[23]   Finkelstein J, Friedman R. Potential Role of Telecommunication Technologies in the Management of Chronic Health Conditions. Dis Manage Health Outcomes 2000;8: 57-63.
[24]   Revere D, Dunbar P. Review of Computer-generated Outpatient Health Behavior Interventions: Clinical Encounters "in Absentia". Journal of the American Medical Informatics Association 2001;8: 62-79.
[25]   Corkrey R, Parkinson L. Interactive voice response: review of studies 1989-2000. Behav Res Methods Instrum Comput 2002;34: 342-353.
[26]   Piette J. Interactive voice response systems in the diagnosis and management of chronic disease. Am J Manag Care 2000;6: 817-827.
[27]   Friedman R. Automated telephone conversations to asses health behavior and deliver behavioral interventions. Journal of Medical Systems 1998;22: 95-102.
[28]   Vinson C, Bickmore T, Farrell D, Campbell M, Saunders E, Nowak M, Fowler B, Shaikh A. Adapting Research-tested Computerized Tailored Interventions for Implementation in Practice. Translational Behavioral Medicine: Practice, Policy, Research. to appear.
[29]   NIH National Cancer Institute. Workgroup for Computerized Tailored Interventions. Providence, RI; 2007.
[30]   NIH Consumer Health Informatics Summit, Workshop on Diffusing technology-based health interventions. Potomac, MD; 2009.
[31]   Musen MA. Modern architectures for intelligent systems: reusable ontologies and problem-solving methods. In: AMIA Symposium; 1998. p. 46-52.

[32]   Lenert L, Norman GJ, Mailhot M, Patrick K. A framework for modeling health behavior protocols and their linkage to behavioral theory. J Biomedical Informatics 2005;38: 270-280.

[33]   U.S. Department of Health and Human Services. Physical activity and health: a report of the Surgeon General. In. Atlanta, GA: U.S. Department of Health and Human Services, Centers for Disease Control and Prevention, National Center for Chronic Disease Prevention and Health Promotion; 1996.

[34]   U.S. Department of Health and Human Services. Healthy People 2010: Understanding and improving health. Physical activity and fitness, 2nd edition. In. Washington, DC: U.S. Government Printing Office; 2000.

[35]   National Center for Chronic Disease Prevention & Health Promotion. BRFSS trends data nationwide grouped by age, for no leisure time physical activity. In: Centers for Disease Control and Prevention; 2005.

[36]   Hawkins RP, Kreuter M, Resnicow K, Fishbein M, Dijkstra A. Understanding tailoring in communicating about health. Health Educ. Res. 2008;23: 454-466.

[37]   Noar SM, Benac C, Harris M. Does tailoring matter? Meta-analytic review of tailored print health behavior change interventions. Psychological Bulletin 2007;133: 673-93.

[38]   OWL Web Ontology Language Overview Available at http://www.w3.org/TR/owl-features/

[39]   Resource Description Framework (RDF): Concepts and Abstract Syntax. Available at http://www.w3.org/TR/2004/REC-rdf-concepts-20040210/

[40]   Task Model Description (CE TASK 1.0)   Available at http://www.ce.org/cea-2018/

[41]   Boxwala A, Peleg M, Tu S, Ogunyemi O, Zeng Q, Wang D, Patel V, Greenes R, Shortliffe E. GLIF3: a representation format for sharable computer-interpretable clinical practice guidelines. J Biomedical Informatics 2004;37: 147-61.

[42]   Hoekstra R. Ontology Representation: Design Patterns and Ontologies that Make Sense. Amsterdam: IOS Press; 2009.

[43]   van Heijst G, Schreiber AT, Wielinga B. Using explicit ontologies for kbs development. International Journal of Human-Computer Studies 1997;46: 183-292.

[44]   Glanz K, Lewis F, Rimer B. Health Behavior and Health Education: Theory, Research, and Practice. San Francisco, CA: Jossey-Bass; 1997.

[45]   Bickmore T, Gruber A, Picard R. Establishing the computer-patient working alliance in automated health behavior change interventions. Patient Educ Couns 2005;59: 21-30.

[46]   Bickmore T, Pfeifer L. Relational Agents for Antipsychotic Medication Adherence In: CHI'08 Workshop on Technology in Mental Health. Florence, Italy; 2008. p. 1-7.

[47]   Niles I, Pease A. Towards a standard upper ontology. In: Proceedings of the 2nd International Conference of Formal Ontology in Information Systems (FOIS-2001). Ogunquit, Maine; 2001.

[48]   OWL-Time Home page. Available at http://www.isi.edu/~hobbs/owl-time.html

[49]   Protege Home Page. Available at http://protege.stanford.edu/

[50]   Rich C. Building Task-Based User Interfaces With ANSI/CEA-2018. IEEE Computer 2009; 42: 20-27.

[51]   Bickmore T, Schulman D, Shaw G. DTask & LiteBody: Open Source, Standards-based Tools for Building Web-deployed Embodied Conversational Agents. In: Intelligent Virtual Agents. Amsterdam, Netherlands; 2009.

[52]   Bickmore T, Picard R. Establishing and Maintaining Long-Term Human-Computer Relationships. ACM Transactions on Computer Human Interaction 2005;12: 293-327.

[53]   Bickmore T, Pfeifer L, Jack BW. Taking the Time to Care: Empowering Low Health Literacy Hospital Patients with Virtual Nurse Agents In: Proceedings of the ACM SIGCHI Conference on Human Factors in Computing Systems (CHI). Boston, MA; 2009.

[54]   Grosz B, Kraus S. Collaborative plans for group activities. In: Proc. 13th Int. Joint Conf. Artificial Intelligence. Chambery, France; 1993. p. 367--373.

[55]   Lochbaum K. A Collaborative Planning Model of Intentional Structure. Computational Linguistics 1998;24: 525-572.

[56]   ECMAScript. Available at http://www.ecmascript.org/

[57]   Sacks H, Schegloff EA, Jefferson G. A Simplest Systematics for the Organization of Turn-Taking for Conversation. Language 1974;50: 696-735.

[58]   Reiter E, Dale R. Building Natural Language Generation Systems. Cambridge: Cambridge University Press; 2000.

[59]   Kopp S, Krenn B, Marsella S, Marshall AN, Pelachaud C, Pirker H, Thórisson K, Vilhjálmsson H. Towards a Common Framework for Multimodal Generation: The Behavior Markup Language. In: Intelligent Virtual Agents. Marina Del Rey, CA; 2006.

[60]   Rich C, Sidner CL, Lesh N. COLLAGEN: Applying Collaborative Discourse Theory to Human-Computer Interaction. AI Magazine 2001.

[61]   Prochaska J, Velicer W. The Transtheoretical Model of health behavior change. American Journal of Health Promotion 1997;12: 38-48.

[62]   Schegloff EA, Sacks H. Opening up closings. Semiotica 1973;8: 289-327.

[63]   Farzanfar R, Frishkopf S, Migneault J, Friedman R. Telephone-linked care for physical activity: a qualitative evaluation of the use patterns of an information technology program for patients. Journal of Biomedical Informatics 2005;38: 220-228.

[64]   Moyers TB, Martin T, Manuel JK, Miller WR. Assessing competence in the use of motivational interviewing Journal of Substance Abuse Treatment 2005;28: 19-26

[65]   Nielsen J. Usability Engineering. San Francisco: Morgan Kaufmann; 1993.

[66]   Computational Models of Health Behavior Change Dialogue. Available at http://relationalagents.com/litebody.html

Appendix - Transcripts of Sample Counseling Sessions

The following is a partial transcript of a sequence of 3 exercise promotion dialogues with a single patient, gathered during the evaluation study. These transcripts were among those evaluated by the expert counselor.

First Conversation

1-1.A: To talk to me, just touch one of the options on your right.
1-2.U: OK
1-3.A: Hi Elissa. My name is Karen. Nice to meet you.
1-4.U: Hi Karen. Nice to meet you.
1-5.A: Great! I am excited about working with you to improve your health.
1-6.U: Me too.
1-7.A: I am going to be your health advisor. My job is to help you set health goals and overcome obstacles, and I will also be giving you lots of tips.
1-8.U: OK.
1-9.A: You need to keep in mind that I am just a computer character with limited capabilities, so I hope you will understand if there are some things I cannot help you with.
1-10.U: OK.
1-11.A: So, how do you feel about talking to a computer character?
1-12.U: It's kind of neat.
1-13.A: That is good to hear.
1-14.A: So, I would like to chat with you every day over the next month to see how you are doing.
1-15.U: OK.
1-16.A: Each day I will ask you some questions about your health.
1-17.U: OK.
1-18.A: I hope it does not become too repetitive, but it is important that I understand how your feelings about health change over time.
1-19.U: No problem.
1-20.A: Do you currently exercise at least five days per week, for thirty minutes or more?
1-21.U: No
1-22.A: During the next thirty days, do you plan to begin exercising five or more days a week, for at least thirty minutes each time?
1-23.U: Yes!
1-24.A: Cool. We'll talk more about exercise the next time we chat.
1-25.U: OK
1-26.A: Well, that's everything for today.
1-27.A: So, when do you want to get together next?
1-28.U: Tomorrow
1-29.A: Okay, take care.

1-30.A: Bye Elissa.
1-31.U: Bye Karen.

Second Conversation

2-1.A: Hi Elissa. Nice to see you again!
2-2.U: Hi, Karen. Nice to see you too.
2-3.A: How are you?
2-4.U: I'm doing okay.
2-5.A: Good.
2-6.A: So, Do you have anything fun planned for the weekend?
2-7.U: No, not really. And you?
2-8.A: I don't have any special plans.
2-9.A: So, how's the weather out there today?
2-10.U: It's pretty good.
2-11.A: Great! I hope you can get outside and enjoy the weather while doing some nice, refreshing exercise.
2-12.U: Ok
2-13.A: So, Elissa, I would like to ask you a few questions about exercise.
2-14.U: Sure.
2-15.A: How motivated are you to exercise, on a scale from 1 to 10?
2-16.U: 4 or 5
2-17.A: How confident are you that you can exercise, on a scale from 1 to 10?
2-18.U: 2 or 3
2-19.A: How are you feeling about your progress with exercise?
2-20.U: I think I'm having problems.
2-21.A: So when I asked you how important exercise is to you, you gave it about 5 out of 10.
2-22.A: What do you think is important about exercise that made you say 5 instead of zero?
2-23.U: If I exercised, things would be better.
2-24.A: Can you tell me what you think would be better if you started an exercise program?
2-25.U: Maybe I would lose weight.
2-26.A: So if you started an exercise program, you would lose weight.
2-27.U: I could also improve my health.
2-28.A: So if you exercised more, your health would be better.
2-29.U: Yes.
2-30.A: Okay. Let me just summarize what we've been talking about.
2-31.U: OK.
2-32.A: If you did start exercising, it might improve your health.
2-33.U: Yes.
2-34.A: If you started exercising, you could lose weight.
2-35.U: I might.
2-36.A: Okay. Thank you for telling me your thoughts about exercise.
2-37.U: You're welcome.

2-38.A: So starting today, we will make a goal to get a little more exercise every day.
2-39.U: Okay
2-40.A: So, before the next time we talk, would you be willing to try and do some walking?
2-41.U: How much should I do?
2-42.A: You don't have to walk fast or for a long distance. The idea right now is just to get moving a little.
2-43.U: Okay, I can do that.
2-44.A: I think we are going to do some great work together.
2-45.A: I am going to assign you some exercise homework to do. Just something simple to help you start thinking about different things you can do to help you achieve your exercise goals.
2-46.U: Sure
2-47.A: When you don't want to exercise, think about what you would normally do. If that action is anything but exercise, try substituting a new action in place of the old one.
2-48.U: Okay.
2-49.A: If you feel tired, you could try going for a walk. It can be energizing and very refreshing.
2-50.U: Okay.
2-51.A: So Elissa, how about trying this for next time?
2-52.U: OK
2-53.A: Great, I think this will help you explore exercise in a different way and will help you to increase your step count a little faster.
2-54.U: Okay.
2-55.A: Great! We've talked about everything for today.
2-56.A: So, I hope we can see each other soon. When can we chat again?
2-57.U: Tomorrow
2-58.A: Okay then, I will be here. Just come by any time.
2-59.A: Goodbye Elissa.
2-60.U: Bye Karen.

Third Conversation

3-1.A: Hi Elissa. Nice to see you again!
3-2.U: Hi, Karen. Nice to see you too.
3-3.A: How is it going?
3-4.U: Good. How are you?
3-5.A: Same as usual.
3-6.A: So, How is your day going? I hope it is going well.
3-7.U: Sorry, I didn't quite catch that.
3-8.A: So, How is your day going? I hope it is going well.
3-9.U: Yes!
3-10.A: Really? What was it like?
3-11.U: Something fun..
3-12.A: Cool.
3-13.A: So, how's the weather out there today?

3-14.U: It's nice out.

3-15.A: Great! I hope you can get outside and enjoy the weather while doing some nice, refreshing exercise.

3-16.U: Ok

3-17.A: Last time I suggested that you think about substituting healthy actions for unhealthy ones. Did you do this?

3-18.U: Yes

3-19.A: Was this assignment useful in thinking about exercise in your life?

3-20.U: Yes

3-21.A: Good, you are making real progress!

3-22.A: Last time, we talked about trying to do a little walking.

3-23.U: Yes

3-24.A: Did you have a chance to go walking since our last conversation?

3-25.U: No

3-26.A: Sorry to hear that. I'm sure you'll manage it next time.

3-27.A: OK, Elissa, can we talk a little about exercise?

3-28.U: Sure.

3-29.A: How motivated are you to exercise, on a scale from 1 to 10?

3-30.U: 4 or 5

3-31.A: How confident are you that you can exercise, on a scale from 1 to 10?

3-32.U: 2 or 3

3-33.A: Are you feeling good about exercise?

3-34.U: I'd like to talk about it.

3-35.A: Is there anything that worries you about your exercise or starting to exercise?

3-36.U: It's easier if I don't exercise.

3-37.A: Tell me more. What do you like about your current lifestyle, without a regular exercise program?

…