Design and Evaluation of a Peripheral Robotic Conversation Companion

Guy Hoffman[1], Oren Zuckerman[1], Gilad Hirschberger[2], Michal Luria[1], Tal Shani-Sherman[2]

[1]Media Innovation Lab
IDC Herzliya
Herzliya 46150, Israel
{hoffman,orenz,michal.luria}@idc.ac.il

[2]School of Psychology
IDC Herzliya
Herzliya 46150, Israel
{ghirschberger,talshani}@idc.ac.il

ABSTRACT

We present the design, implementation, and evaluation of a peripheral empathy-evoking robotic conversation companion, *Kip1*. The robot's function is to increase people's awareness to the effect of their behavior towards others, potentially leading to behavior change. Specifically, Kip1 is designed to promote non-aggressive conversation between people. It monitors the conversation's nonverbal aspects and maintains an emotional model of its reaction to the conversation. If the conversation seems calm, Kip1 responds by a gesture designed to communicate curious interest. If the conversation seems aggressive, Kip1 responds by a gesture designed to communicate fear. We describe the design process of Kip1, guided by the principles of *peripheral* and *evocative*. We detail its hardware and software systems, and a study evaluating the effects of the robot's autonomous behavior on couples' conversations. We find support for our design goals. A conversation companion reacting to the conversation led to more gaze attention, but not more verbal distraction, compared to a robot that moves but does not react to the conversation. This suggests that robotic devices could be designed as companions to human-human interaction without compromising the natural communication flow between people. Participants also rated the reacting robot as having significantly more social human character traits and as being significantly more similar to them. This points to the robot's potential to elicit people's empathy.

Categories and Subject Descriptors

H.1.2 [**Models and Principles**]: User/Machine Systems; J.4 [**Computer Applications**]: Social and Behavioral Sciences—psychology.

General Terms

Experimentation, Human Factors.

Keywords

Human-robot interaction; Design; Robotic companions; Behavior change; Empathy; Ambient kinetic tangibles; Smartphone robots.
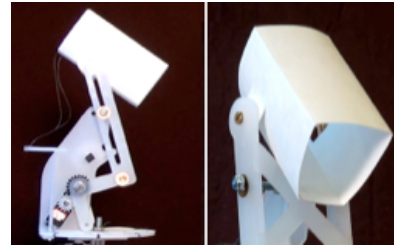
Fig. 1. *Kip1*, a peripheral robotic conversation companion.

1. INTRODUCTION

When people interact, they are often unaware of the effect their behavior has on others. To address this issue we are designing and developing a series of peripheral robotic companions, aimed to accompany natural human-human interaction, and reflect the effect of one's behavior through subtle physical gestures. We hope that the presence of such a peripheral companion may lead to increased awareness among the interacting humans without compromising their natural communication pattern. This paper describes such a robot, *Kip1*, designed to promote non-aggressive conversation between people.

Kip1 is designed as a small desktop structure, reminiscent of a lamp (Fig. 1). When a conversation is taking place near Kip1, it monitors the nonverbal content of the conversation, e.g., speech timing, silences, and loudness. The robot tracks speaking vs. silent segments, and the ongoing and incidental loudness of the conversants. If there is no ongoing conversation, Kip1 is in a calm, relaxed state, indicated by a slow, deep "breathing" gesture. If an ongoing conversation is calm, Kip1 shows interest by stretching upwards in a "curious" gesture towards one of the participants. If, however, the conversation becomes too loud, Kip1 retracts into a "scared" gesture, shivering and lowering its head.

Unlike most human-robot interaction research, concerned with direct interaction between people and robots, this project's design goal is to supplement face-to-face human-human interaction. Peripheral conversation companions are meant to influence and enhance direct human interaction, rather than replace it, mediate it, or distract from it.

In this paper, we present Kip1's design process, including the design considerations and choices made with regards to the robot's material, mechanism, gestures, software, and hardware. We had two design goals: A balance between drawing people's attention and not distracting them from their current conversation; and a design eliciting an emotional connection between people and the robotic device.
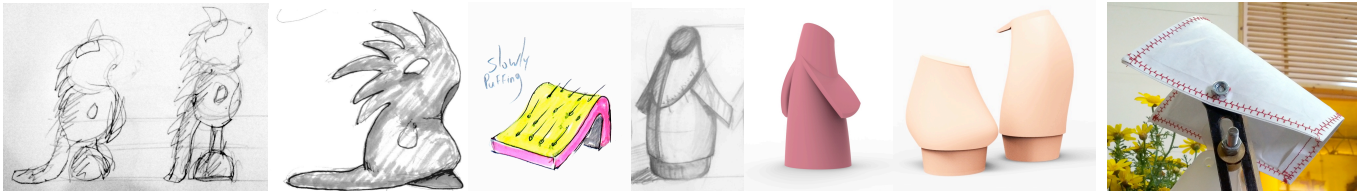
Fig. 2. Early pencil sketches (left); 3D renders of winged blob design (middle); Early head designs with heavy stitching (right).

We evaluated our design in an experimental setting. Participating couples were guided to find a topic of severe disagreement and talk about it with their partner while the robot was in the room. The experimental group shared the room with an autonomously reacting robot and the control group with an animated, but not-reacting robot. We used both quantitative and qualitative methods to analyze the interaction. After reporting on our findings, we relate them to our design guidelines and considerations, and conclude with future work.

2. RELATED WORK

Kip1 can be thought of simultaneously as a socially expressive robot, an ambient kinetic tangible, and a conversation monitoring interface. We survey related work from all three domains, as well as from work related to couples in conflict.

2.1 Robotic Nonverbal Expressions of Affect

Socially interactive robots use both verbal and nonverbal channels in order to express their emotional state. In fact, Fong *et al.* describe the capability to express emotions as one of the indicators of socially interactive robots [5]. In anthropomorphic robots, facial expressions are often used to express emotions, either on a screen [7, 20] or using actuated facial features [1, 19]. Robots that do not have an expressive face or are non-anthropomorphic at all can use gestures to express emotions [2, 13]. In some cases, for robots that have no social articulation at all, such as UAVs, path planning has been used to express emotions [22]. Virtually all of these systems are used either for direct human-robot interaction, or for performance robotics. Our approach differs in that we use the nonverbal affective expression as an ambient companion to human-human interaction.

2.2 Technology mediated conversation

Prior technologies that mediate human conversation are usually screen based. DiMicco *et al.* used a shared display in a group interaction, showing how much each participant contributed to the conversation [4]. A similar study by Bergstrom & Karahalios used a "conversation clock" screen that visualized the time each participant talked, but without associating the visualization with a specific participant [3]. In contrast, our system uses ambient physical gestures, a tangible, embodied modality.

2.3 Ambient Kinetic Tangibles

This project lies at the crossroads of HRI and Tangible User Interfaces (TUI) research. In particular, Kip1 exemplifies the following core aspects of TUI: (a) "coupling digital information to everyday physical objects and environments" [15], and (b) "providing tangible representation to digital information" [21]. The robot's physical gestures are tangible representations of its emotional model, which is the digital information reflecting the conversation happening around the device. Kip1 also follows the "Objects for Change" [24] principle of implementing established behavior change techniques in the design of a TUI device.

Ambient interfaces use visual and auditory cues designed to be processed at the periphery or background. Usually they use subtle changes in light and sound to represent digital information [15]. In some cases ambient interfaces use tangible representation, mapping digital information to physical motion, thus being a "kinetic tangible" [14]. Kip1 continues this tradition, but in addition monitors real-time local information.

2.4 Vocalics in Relationship Conflict

Research has consistently documented how the emotional climate of relationship conflict interactions is an important marker of overall relationship quality. Emotional exchanges characterized by high levels of negative emotional behavior and low levels of positive emotional behavior have been associated with greater marital dissatisfaction and instability [9, 16]. An important indicator of emotion in marital interactions is emotional verbal and nonverbal behavior. Previous studies show that vocal expression demonstrates a variety of emotions that can eventually affect relationship quality and happiness. Shaver *et al.*, for example, found that loud voice, yelling and screaming is perceived as an expression of anger [23]. Kip1 intends to address this issue by using socially expressive robotics in the context of autonomous conversation monitoring.

3. DESIGN PROCESS

In the following section we present the Kip1 design process, including physical appearance, material selection, DoF layout, and gesture design. We describe our design guidelines, the interdisciplinary design team, and the various stages of design prototyping up until the current prototype, the third in sequence.

3.1 Process Overview

We designed Kip1 using a combined interaction, animation, and industrial design process, similar to the one proposed by Hoffman & Ju [13]. In the process, we explored a large number of forms and gesture capabilities, through a variety of techniques: pencil sketches; animation studies; mechanical CAD designs; skeleton prototypes; material explorations; and three actuated increasingly finalized physical prototypes.

3.2 Design Guidelines

Kip1's pronounced goal is to accompany human-human interaction, peripherally tracking the aggressiveness of their conversation. Through its emotional response, the robot should promote the conversants' awareness to the possible emotional effect of their behavior, hopefully leading to behavior change. We thus defined two design guidelines as a goal for our process:

3.2.1 Peripheral

The robot is meant to work in the periphery of the human-human interaction. No interaction is intended to occur directly between the human and the robot. Instead, it should be perceived in an ambient fashion. This is particularly important, since our goal is to promote and supplement face-to-face human-human interaction, to slightly influence and enhance it, but not replace it or distract
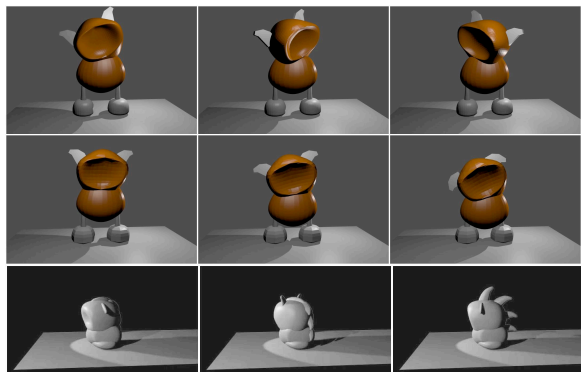
Fig. 3. Still frames from 3D animation sketches.

Fig. 4. Various sketches of multi-action linkage allowing two DoFs (rising and expansion) to be driven by a single motor.

from it. That said, the robot's effect should be successful even without full attention from the human conversants.

3.2.2 Evocative
The robot's appearance and behavior should evoke empathy on the side of the human. Ideally, the human raising their voice and experiencing Kip1's fear gesture, should feel slightly bad for the robot's perceived "hurt feelings", and as a result consider the effects of their behavior on others.

One of the ways Kip1 would be able to evoke empathy was by seeming fragile itself. The robot's fragility could help highlight the potential sensitivity of the conversation partner which they themselves were unable to express. Therefore, instead of suggesting strength, stability, and efficiency, traits often associated with robot design, we wanted Kip1 to communicate fragility and sensitivity. In particular, we wanted these to be expressed in all of the robot's structure, materials, and movement.

3.3 Interdisciplinary Design Team
Robot design is often driven by engineering requirements and later "finished" with industrial design. Since we are specifically interested in robot design as a practice, one of the explicit project goals was to involve an interdisciplinary team throughout. Viewing this project as lying on the intersection of HRI and TUI research (see: Section 2), the team included academic researchers from both fields. To balance the engineering and technical development with a design-oriented focus, we included an industrial designer and a puppet designer from the beginning of the project, in addition to a mechanical engineer, a hardware prototyping expert, and a computer science student.

The design guidelines were shared with the various team members, and the iterative process allowed each member of the team to contribute using their field of expertise and influence the whole team. For example, the puppet designer addressed the "fragile" design guideline with several variations for the robot's head made from soft materials, from cloth to paper. The mechanical engineer was influenced by the use of soft materials and was inspired to add evocative motion to the robot's head using a piece of string instead of adding another motor.

3.4 Pencil Sketches
We started the design process with iterative rounds of pencil sketches and animation studies. These suggested a number of simple forms and motion axes, and can be thought of as very free-form improvisations on the theme of a sensitive, evocative creature. We posited that a small animal-like form and behavior would be most appropriate, as humans are extremely successful at
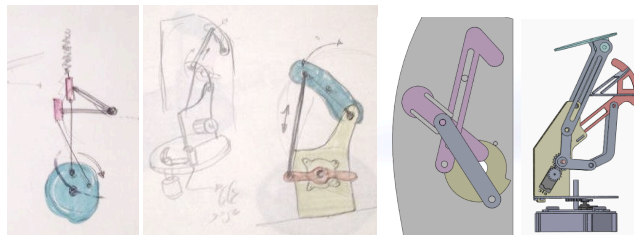
reading (perhaps, into) animals' emotional states. We also know from pets that animals often successfully evoke empathy in humans. Fig. 2 (left) shows a selection of sketches from this stage.

We went back and forth between the paper-based sketches and simple, quick-and-dirty animation studies, each informing the other. Small animal-like and creature-like forms emerged with each of them having a "curious" or "confident" state and a "fearful" state. The first one was usually indicated by expansion or rising, and in some cases by the addition of ear-perking. The second was usually indicated by contraction, and sometimes by the addition of hair or spikes rising out of the body of the creature. This stage of the design increasingly moved towards more abstract and non-anthropomorphic shapes.

3.5 Animation Studies
As suggested by Hoffman & Ju [13], we explored the relationship between robot appearance and movement in a series of 3D animation studies. Fig. 3 shows still frames from these studies. We were mostly interested how clearly the robot's emotional state might read, in particular with a limitation of few degrees of freedom (DoFs). It is important to stress that each stage of animation study led to another round of drawing sketches towards the next animation study. We also treated these studies as very rough sketches, each made in less than an hour, and without expressively delving into detailed features of the robot's design. They can be thought of as "mass studies", where abstract shapes move with respect to one another.

As we moved away from anthropomorphic shapes, we remained in the realm of organic motion, perhaps like that found in fish or invertebrates. We thought that using an abstract shape with organic seeming motion would successfully bridge the tradeoff between mechanical feasibility and easy-to-read gestures.

Our animation tests suggested the following movement paradigms to support our usage and interaction scenarios:

Rising and falling — Growing and shrinking vertically *in size* resulted in a readable indicator of confidence and curiosity.

Inflating and deflating — Growing and shrinking *in volume* read very effectively, and seemed to support the idea of both confidence and fear, even in peripheral vision, as humans are acutely aware of changes in size.

Shivering — We found shivering to portray fear in an unambiguous manner.

Protruding spikes — We explored the idea of protruding spikes as an indication of extreme fear (Fig. 2 middle and Fig. 3 bottom). However, for mechanical reasons, they were not included in the current prototype. We hope to include them in future revisions.

At this stage, the team converged on a design that was centered around a simple, organic shape, akin to an egg or a penguin (Fig. 2 middle), with the following gestures: Rising and expanding to express self-confidence; contracting, falling, and shivering to express fear. In addition, we planned to have left and right movement to indicate attention, and rising and falling wing-like shapes to support additional gestures and emotional states.

3.6 Material Exploration

We wanted the materials to also reflect the peripheral nature of the robot, as well as its fragility. In addition, we were keen to evaluate the use of alternative materials in the design of robots, beyond the classic aluminum, steel, and plastics so often utilized. To that end, we collaborated with a puppet designer to explore the integration of different materials with moving parts.

We worked with various kinds of cloths, from thin silk-like fabric to heavy denims of various colors. We tested the interrelation between the robot's movement and the flexibility of the fabric. We explored wood in various forms, such as solid, plywood, and fiberboard. We also tested non-structural metals, such as copper and brass. Finally, we tested the use of paper as both a structural and aesthetic element in the robot's design. In particular, we experimented with Tyvek™, a non-woven synthetic paper-like material, that can be folded into shape, does not rip, and has a glossy surface.

At some stage, the focus of the design moved to stitching. Both paper and fabric can be stitched, and the exposed stitches can communicate fragility and emphasize the robot's form. We tested bold, thick stitching patterns on the robot's body (Fig. 2 right), but eventually abandoned that particular design path. Our final design was made up of a heavy MDF base, with light-colored translucent acrylic joined by brass joints, and white Tyvek.

3.7 Mechanism Studies

The material exploration went hand-in-hand with mechanism designs, which were intended to go into the fabric or paper shell. These designs were done both in 3D CAD software, and with low-fidelity cardboard prototypes.

3.7.1 Multi-Action Linkages

As we were interested in the expressive potential of a low-DoF robot, we recruited two mechanical principles: Multiple-action linkages and secondary action. Multiple-action linkages allow for more than one movement using the control of a single motor. In particular, one DoF can control several distinctly moving parts, and can cause different directions of movement throughout the trajectory of a single motor (Fig. 4).

Multi-action linkages are both economically efficient and mechanically elegant. That said, one of the drawback of multi-action linkages is that they are more difficult to model mathematically, and therefore more difficult to control. However, if we map the various motions onto a single driving variable, this issue can actually be avoided. In our case, the "confidence" variable maps directly onto one motor, which then causes an effect in several confidence-related movements, such as expansion and rising, circumventing the control problem.

3.7.2 Bare Mechanism

We built a first physically actuated prototype of the driving mechanism, while constructing a fabric shell. This was to better understand the physical constraints and dynamic properties of the robot's structure, and to kick off electronics, control, and software

design. At this stage, however, we found that the exposed mechanism provided for a more delicate and fragile appearance than when it was covered in a fabric or paper shell. The weight of the additional material made the robot seem more clumsy, heavy, and solid, whereas the bare mechanism seemed exposed and sensitive. We thus proceeded to explore the idea of using only the mechanism as the robot's appearance design. In addition, we believed that the tension between the robot's mechanical appearance and organic movement would balance well.

3.7.3 Secondary Action

Secondary action is an animation principle which can improve the motion characteristic of a robot by adding a passive DoF that is influenced by the actuated movement and physical constraints, including gravity, enriching the dynamic perception of the robot [6]. Since one of the gesture primitives we designed was a shiver, we thought that the amplification of this gesture by a secondary DoF would be effective. We thus added a passive DoF, in the form of a loose head-like shape to the robot's structure.

Combining the principles of multi-action linkage and secondary action, we tied a thin thread between the back side of the robot's neck and its body. This had two positive effects: It added an additional movement at the end of the motor's trajectory caused by the tensioning of the string. The string also communicates fragility as it is clear that breaking it would sever the naturalistic connection between the robot's head and body.

3.8 No Screens

In order to maintain the focus of Kip1's users on each other, it was important in our design process to refrain from using screens as part of the interaction paradigm. Some other recent desktop robots use mobile devices as their sensor and processing platform [11, 17]. Usually the screen is used for expressive face-like features and animations, or to display text. We made the design decision to express all feedback through physical gestures alone. This was based on the consideration that to support direct human-human interaction, physical gestures are less distracting than screens, and that if our aim is a gentle nudge towards behavior change, gestures can play a more subtle role than on-screen information. Moreover, as Kip1 is supposed to be in the background, embodied spatial movement is more easily read in peripheral view than on-screen feedback.

4. HARDWARE AND SOFTWARE

The robot's system design uses a smartphone as the main sensing and computing hardware [11], and includes four main components: An Android smartphone running the sensing and control software of the robot, a IOIO microcontroller board linking the smartphone to the motors, two servo motors, and a mechanical structure using a variety of linkages to express the robot's gestures. The smartphone uses its internal microphone to monitor the ongoing conversation around it, and runs a single application, consisting of four modules: Volume Detection, Conversation Analysis, Emotional Model, and Gesture Controller. The Volume Detection module constantly records real-time audio and measures the current volume of the audio coming in. It then compares this audio with a baseline room-level and outputs a relative volume associated with the conversation. This value is sent to the Conversation Analysis and Emotional Model modules.

4.1 Conversation Analysis

This module maintains a conversation state using a ring array of detection windows and a finite state machine (Fig. 5a). For each i
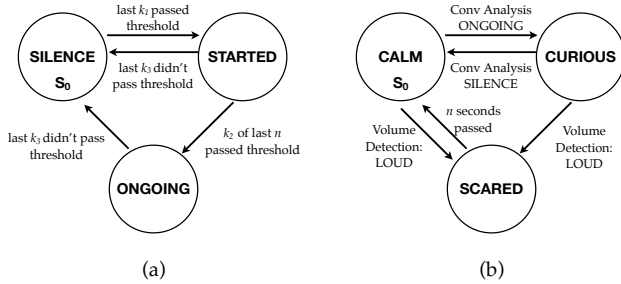
Fig. 5. Kip1 Finite State Machines for (a) Conversation Analysis and (b) Emotional Model.

seconds of audio, the module tags the next node in a ring buffer of size n with the maximum relative volume for that time period. If the last k_1 windows all surpassed the conversation level threshold, it is considered a conversation (as opposed to stray noise). This puts the state machine into the STARTED state. If k_2 of the last n windows surpass the conversation level threshold, this is considered an ongoing conversation, and the state machine switches to ONGOING. k_3 windows below threshold volume switch the state machine back to SILENCE. In our implementation we experimented with several values, and found $i=2, n=10, k = \{3,10,3\}$ to be most successful.

4.2 Emotional Model

The Emotional Model monitors both the conversation state and the real-time incidental volume and is also implemented as a finite state machine (Fig. 5b) in combination with a continuous "self-confidence" variable CONF between 0 and 1 (not shown), initialized to 0. The CONF value rises as long as the Emotional Model is in the CALM state. If the conversation state moves from STARTED to ONGOING, the emotional state moves from CALM to CURIOUS. If, at any point, a loud noise is detected, CONF resets to zero, and either state moves to SCARED, which is maintained until a pre-set timer elapses (currently: 5 seconds). The Emotional Model then returns to a CALM state.

4.3 Gesture Controller

The Gesture Controller monitors the Emotional Model and runs appropriate gesture behaviors based on state transitions in the Model. This module is based on the Gesture System in [11]. At any given moment, it runs one of a set of Behaviors. Some behaviors are one-offs, and some are cyclical. In our implementation, one-off behaviors resolve into repetitive behaviors automatically to prevent the robot from "freezing". For example, in the CALM state, the Behavior Controller runs a looping movement akin to slow, deep breathing, around the lower edge of the robot's movement. The amplitude of the breathing is affected by the Emotional Model's self-confidence value. In the CURIOUS state, the robot stretches out towards the conversant, looks around, and, using a string as a mechanical linkage, raises its "head" upwards. In the SCARED state, the robot retracts to a fully cowering state, and shivers.

The Gesture Controller includes a Behavior Controller, managing the behavioral state, a Trajectory Interpolator, making sure the movements are smooth and appealing according to character animation principles, and a Motor Controller which manages a model of each DoF, motor positions and limits, and its role in the mechanical structure. This system then uses the built-in IOIO messaging system to set motor positions at a fixed frequency (currently: 200Hz).

5. EVALUATION

We performed an evaluation of our design in the context of conflict conversation between couples. We wanted to see how the robot's design and behavior—i.e. its reaction to the couple's vocalics—might influence the conversation. We were also interested in the sentiments and opinions the robot evokes. We were specifically looking for findings that related to our two core design goals: a *peripheral* and *evocative* robotic companion.

To do so, we conducted quantitative and qualitative analysis of a controlled laboratory experiment, in which participant couples were guided to find a topic of disagreement, and asked to talk about this topic for 15 minutes. The participants were told that we were testing a new robot device which will listen in on their conversation. They either shared the room with a reacting robot or with a non-reacting, but alive-seeming robot. We recorded the couple's conversation with video and audio. In the end, we asked couples to fill out a brief questionnaire about their experience.

5.1 Participants

A total of 30 heterosexual Israeli couples (60 participants total) participated in the experiment. Participants were recruited from social networks in return for the equivalent of roughly $41 USD at time of writing. We recruited couples whose relationship was at least six months long. Participants' age ranged from 21 to 30 (M = 24.64, SD = 1.83).

5.2 Method

We conducted a single independent-variable between-subject (between-couples) design. We manipulated one variable, the robot's physical reaction to the conversation. In the EXPERIMENT condition, the robot responded with physical gestures to its emotional state as described above. In the CONTROL condition, the robot monitored the conversation, switched emotional states internally, but did not move in response to these states. Instead, it maintained the same gentle breathing behavior as in the CALM state of the EXPERIMENT condition.

5.3 Procedure

The experiment was conducted in experiment rooms with controlled lighting, no windows, and no outside distractions. Upon arrival, each participant couple was welcomed into the first room, where the experimenter explained the initial experimental guidelines and obtained informed consent. Then, the couple was split up, with one participant asked to enter the second experiment room. In separate rooms, couples completed the Couple's Problem Inventory [8], in which they rated the perceived severity of disagreement of a standard set of marital issues such as money, in-laws, and sex. The experimenter then helped the couple select an issue, which both spouses rated as being of high disagreement severity, to use as the topic for the conflict conversation.

Next, participants were guided to the third, adjacent, room, were asked to sit on two facing chairs, with the robot sitting on a coffee table between them and slightly offset to the side. The robot, measuring about 30cm in height, reached approximately to the humans' mid-torso when they were seated. Based on Leveson and Gottman [18], the couple was asked to discuss the chosen topic for fifteen minutes. They were told that we are testing a new robotic device which will listen in on their conversation and "may or may not react to it". After indicating that they understood these instructions, the participants were left alone in the room to talk about their topic.
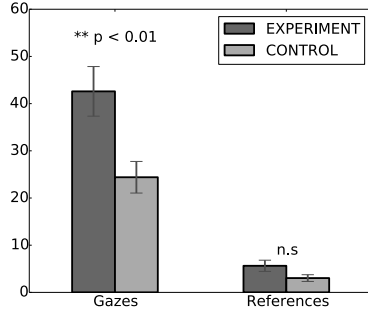
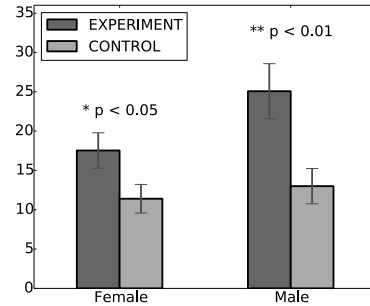**Fig. 7. Gaze towards robot and verbal references to it by condition.**

Fig. 8. Gender differences in number of gazes towards the robot.
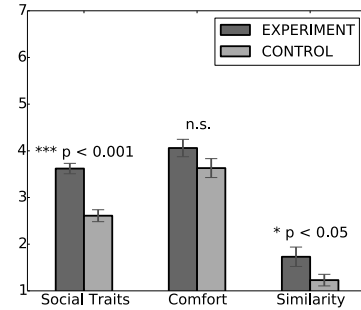

Fig. 9. Subjective measures. All variables were on a scale of 1–7.

Upon completion of the interaction stage, the experimenter re-entered the room, and asked the participants to fill out the post-procedure questionnaires, in separate rooms. This included the self-report measures below and a demographic questionnaire.

Participants were then fully debriefed, and were explicitly told that they were artificially put in a situation of conflict as part of this experiment, and that any negative feelings that might have come up during this conversation are likely to have originated from the experimental setting. We made sure that they felt good about their participation in the study before concluding.

6. MEASURES

We used both quantitative and qualitative methods to analyze the experiment. These measures were chosen to evaluate the design principles laid out in Section 3. For quantitative measures, we measured both objective behavioral measures coded from the experimental video, and subjective metrics from questionnaires:

6.1 Behavioral Measures

For each occurrence of the behavioral measures, we used the mean count of two independent coders who were unaware of the experimental hypotheses.

Gaze at Robot— We coded gaze events towards the robot for each member of the couple, as a measure of attention towards the robot.

Verbal References to Robot— We coded verbal references to the robot for each couple, as a measure of the robot disrupting the dyadic interaction.

6.2 Questionnaires

All responses on the questionnaires were on a scale of 1–7.

Robot Social Human Character Traits—Participants rated their impression of the robot on a composite measure of five items indicating positive social human character traits, including the robot's perceived friendliness, confidence, warmth, cooperativeness, and sociability. This measure was validated and found reliable in previous studies [10, 12]. Cronbach's α was good for this scale in this study, too (.78).

Comfort Level with Robot—Participants rated their comfort level with the robot on a composite measure of eight items, including their sense of being understood, being annoyed by the robot (reverse scale), feeling relaxed, connected, comfortable, and silly (reverse scale). Cronbach's α was good for this scale (.77).

Robot Similarity—We asked participants to rate on one scale to what extent the robot was similar to them, as a measure of potential empathy with the robot.

6.3 Qualitative Analysis

We also conducted qualitative analysis of prototypical videos from the EXPERIMENT condition.

7. RESULTS

To recap, participant couples were randomly assigned to hold their conversation in one of two conditions: Couples in the EXPERIMENT condition, conversed in the presence of a robot that reacted to their speech. Couples in the CONTROL condition, conversed in the presence of a robot that displayed a regular breathing pattern, but did not react to their speech.

7.1 Gaze At Robot

Gaze at the robot was gleaned from video coding. An independent means t-test on the number of gazes at the robot yielded significant results, $t(28) = 2.92$, $p < .01$. Couples in the EXPERIMENT condition looked almost twice as often at the robot (M = 42.6, SD = 19.67), than those in the CONTROL condition (M = 24.4, SD = 12.53) (Fig. 7 left).

7.1.1 Gender Differences

Measuring gaze by men and women separately reveals that the difference in gaze behavior was more pronounced in men than in women, albeit significantly different for both (Fig. 9). For men, an independent means t-tests on the number of gazes at the robot yielded significant results, $t(28) = 2.90$, $p < .01$. Men in the EXPERIMENT condition looked almost twice as often at the robot (M = 25.1, SD = 13.11), than those in the CONTROL condition (M = 13.0, SD = 8.39). For women, an independent means t-tests on the number of gazes at the robot yielded significant, but less pronounced results, $t(28) = 2.12$, $p < .05$. Women in the EXPERIMENT condition looked about 50% more often at the robot (M = 17.5, SD = 8.42), than those in the CONTROL condition (M = 11.4, SD = 6.80). Overall, women looked less often at the robot than men, especially in the EXPERIMENT condition.

7.2 Verbal References to the Robot

Verbal references to the robot were also obtained by video coding. An independent means t-test on the number of gazes at the robot yielded no significant results, $t(28) = 1.87$, n.s.. Couples in the EXPERIMENT condition referred to the robot slightly more often (M = 5.63, SD = 4.45) than those in the CONTROL condition (M = 3.03, SD = 2.67). However between-couple variance was high. Moreover, number of mentions were very low compared to the number of gazes towards the robot (Fig. 7 right).

7.3 Robot Social Human Character Traits

This composite variable measured positive social human traits attributed to the robot. An independent means t-test on the average of this scale yielded significant results, $t(28) = 3.55$, $p < 0.001$. Couples in the EXPERIMENT condition rated the robot's social human traits as higher (M = 3.62, SD = 0.61), than those in the CONTROL condition (M = 2.61, SD = 0.70). (Fig. 9 left).

7.4 Comfort Level with Robot

This composite variable measured how comfortable participants were to have the conversation with the robot present. An independent means t-test on the average of this scale yielded no significant results, $t(28) = 1.57$, n.s.. Couples in the EXPERIMENT condition were similarly comfortable with the robot (M = 4.06, SD = 1.02), than those in the CONTROL condition (M = 3.63, SD = 1.11). (Fig. 9 middle)

7.5 Similarity to Robot

This single measure ranked how similar participants viewed the robot to themselves. An independent means t-test on the average of this scale yielded significant results, $t(28) = 2.06$, $p < 0.05$. Couples in the EXPERIMENT condition rated the robot as slightly more similar to them (M = 1.73, SD = 1.14), than those in the CONTROL condition (M = 1.23, SD = 0.68). (Fig. 9 right).

7.6 Qualitative Analysis

We informally explored interaction patterns between participants and the robot by studying the experimental videos. Below is a qualitative analysis of different ways in which couples reacted to the robot, illustrated by two selected couples. To select the couples, we qualitatively evaluated the interaction of all 15 couples in the EXPERIMENT condition, and classified them into types according to their reaction to Kip1. We then chose representative couples for two different types of interactions to be included in this paper. The chosen couples were both engaged in heated conflict, but displayed responses to the robot at both ends of the spectrum. Couple 1 (C1) was relatively reactive to Kip1, while Couple 2 (C2) ignored it for the most part, getting lost in heated conversation.

At the beginning of C1's discussion C1F (Female) was talking calmly to C1M (Male). Kip1 reacted appropriately with a "curious" gesture, in this case towards C1M. C1M briefly looked back at Kip1 with a smile and then shifted his gaze back to his partner. Kip1 then leaned curiously towards C1F, causing her to stop mid-sentence, laugh and say: "It stresses me out! […] he is looking at me!". She then naturally looked back to continue the discussion.

Reactions to the robot differed when the couple was engaged in a heated discussion, showing that Kip1 was not distracting when their attention was fully directed towards their discussion partner. For example, later in the conversation, C1M, who previously reacted, ignored the "curious" gesture when the discussion topic (but not volume) was getting more heated.

In C2, C2M was constantly speaking in an aggressive tone causing Kip1 to continuously react with the "scared" gesture. The couple looked at Kip1 the first time it was scared, but C2M said dismissively: "This is so stupid", and ignored Kip1 from that moment on. Along the conversation, when C2F was speaking softly, Kip1 barely started to shift towards a "curious" gesture but every time C2M interrupted her, the robot returned to "scared", so the couple never saw Kip1 fully displaying its "curious" gesture.

In some cases, people turned to Kip1 when they seemed to be in an uncomfortable situation. C1F made an argument and raised her voice, Kip1 reacted appropriately with a "scared" gesture, but the couple ignored it as they were engaged in their own discussion. A few seconds later, while C1F was still making her rather loud argument and Kip1 was still in a "scared" gesture, C1M shifted his gaze and looked at Kip1. C1F followed C1M's lead and also shifted her gaze toward the "scared" robot, stopped talking, and remarked: "He has an epileptic attack I think". They both laughed shortly, breaking the tension, and allowing C1M to speak.

In C2, where Kip1 was constantly in a "scared" gesture due to C2M's dominating tone, C2F often glanced towards Kip1, especially when C2M was speaking particularly loudly. These glances were usually accompanied by seemingly nervous self-adaptation behavior, such as playing with her fingers or hair, and could be understood as an attempt to shift the focus toward something outside of the uncomfortable conversation.

In some cases, the conversation changed after participants' reaction to Kip1. For example, as mentioned above, C1F and C1M moved their attention from the heated discussion (led by C1F) towards Kip1, with a "comic relief" by C1F about Kip1's shivers. After a few seconds, C1M reinitiated the discussion, but instead of making a counter argument to C1F, he gave a meta-level perspective of the situation, stating that they have different opinions about the topic.

8. DISCUSSION AND FUTURE WORK

Our findings show significant differences between conditions in the number of gazes at the robot, in the robot's perception of positive social character traits, and in the perceived similarity to the robot. We found no significant difference in the number of verbal reference to the robot, and in the comfort level to converse next to the robot. These findings suggest two main insights.

Intriguing but not distracting: Participants directed significantly more attention towards the robot when it reacted to their conversation—but they did so only via gaze, and not verbal reaction, which remained the same in both conditions. This supports our *peripheral* design goal: People react to the robot's behavior in a way that does not interfere with the natural flow of the human-human conversation. A temporary shift of gaze away from a partner is not considered rude or distracting. Our findings thus support the notion that our design was intriguing without being distracting.

An object with social-emotional appeal: When the robot reacted to participants' conversation they perceived it as more human-like and more similar to them. This supports our *evocative* design goal: The robot was perceived as more friendly, warm, social, and similar to the participant, meaning people did not relate to it as an object, but as a device that might be perceived as capable of forming human emotions, bond, or attachment. Participants felt equally comfortable to converse in the presence of the robot in both conditions, suggesting that the robot's gestures, not just appearance caused a difference in perception on the social-emotional level.

Anecdotally, it is worth noting that in a measure-by-measure analysis of the social character scale, the only trait that *wasn't* significantly different between the conditions was the one rating the robot as "confident", suggesting that our design goal of the robot's fragility (Section 3.2.2) transpired.

Still, our study did not specifically check whether couples correctly read the gestures or actually empathized with the robot. We are now running follow-up studies to evaluate these questions.

8.1 Future Work

We plan to further develop the conversation analysis capabilities of the robot in order to pick up on more complex vocalic cues. To do so, we are now working to analyze pitch together with loudness to categorize vocal affect in speakers, with the aim of achieving more precise detection of aggressive speaking behavior as well as more expressive gesture feedback. We are also working on systems to separate speakers and add the relative contribution of each speaker and overall conversation balance to the robot's emotional model. Furthermore, we are experimenting with adding physiological measures, such as heart rate and galvanic skin response to the robot's sensing capabilities.

We plan to extend the robot's potential for social bonding and emotional connection, through new degrees of freedom and gestures, including going back to some abandoned design ideas.

Finally, we plan to study the interpersonal effects of this robotic companion, including the effect it had on the amount of aggressive conversation, and how couples felt about each other after speaking in the presence of the robot. We also want to study our approach in additional setting beyond couples interaction, for example parent-child or teacher-student interaction.

9. CONCLUSION

We presented a design and evaluation of a peripheral robotic conversation companion. Our findings support our hypothesis that a peripheral and evocative robotic design can accompany human-human interaction in an intriguing yet non-distracting way, as an object with perceived social and emotional traits. We conclude that robotic devices could be designed to evoke empathy among people and serve as companions to human-human interaction without compromising the natural communication patterns between humans. We see such conversation companion robots as having applications for conflict mediation, classroom settings, business meetings, dating, negotiations, and more.

10. REFERENCES

[1] Bartneck, C. et al. 2004. In your face, robot! The influence of a character's embodiment on how users perceive its emotional expressions. *Proc of the Design and Emotion*.

[2] Beck, A. et al. 2010. Towards an affect space for robots to display emotional body language. *RO-MAN 2010 - The 19th IEEE International Symposium on Robot and Human Interactive Communication* 464–469.

[3] Bergstrom, T. and Karahalios, K. 2009. Social mirrors as social signals: transforming audio into graphics. *Computer Graphics and Applications, IEEE*. 29, 5, 22–32.

[4] DiMicco, J.M. et al. 2004. Influencing group participation with a shared display. *Proc of the 2004 ACM conference on Computer supported cooperative work* 614–623.

[5] Fong, T. et al. 2003. A survey of socially interactive robots. *Robotics and Autonomous Systems*. 42, 3, 143–166.

[6] Gielniak, M.J. et al. 2010. Secondary action in robot motion. *19th Int'l Symposium in Robot and Human Interactive Communication (RO-MAN 2010)* 310–315.

[7] Gockley, R. et al. 2005. Designing robots for long-term social interaction. *2005 IEEE/RSJ International Conference on Intelligent Robots and Systems.(IROS 2005)*. 1338–1343.

[8] Gottman, J. et al. 1977. The Topography of Marital Conflict: A Sequential Analysis of Verbal and Nonverbal Behavior. *Journal of Marriage and the Family*. 39, 461.

[9] Gottman, J.M. et al. 1998. Predicting Marital Happiness and Stability from Newlywed Interactions. *Journal of Marriage and the Family*. 60, 1, 5.

[10] Guadagno, R.E. and Cialdini, R.B. 2002. Online persuasion: An examination of gender differences in computer-mediated interpersonal influence. *Group Dynamics: Theory, Research, and Practice*. 6, 1, 38–51.

[11] Hoffman, G. 2012. Dumb Robots , Smart Phones: a Case Study of Music Listening Companionship. *RO-MAN 2012 - The IEEE Int'l Symposium on Robot and Human Interactive Communication* 358–363.

[12] Hoffman, G. et al. 2014. Robot responsiveness to human disclosure affects social impression and appeal. *Proceedings of the 2014 ACM/IEEE international conference on Human-robot interaction* 1–8.

[13] Hoffman, G. and Ju, W. 2014. Designing Robots With Movement in Mind. *J of Human-Robot Interact*. 3, 1, 89.

[14] Ishii, H. et al. 2012. Radical atoms: beyond tangible bits, toward transformable materials. *Interactions*. 19, 1, 38–51.

[15] Ishii, H. and Ullmer, B. 1997. Tangible bits: towards seamless interfaces between people, bits and atoms. *Proc of the SIGCHI conference on Human factors in computing systems* 234–241.

[16] Karney, B.R. and Bradbury, T.N. 1997. Neuroticism, marital interaction, and the trajectory of marital satisfaction. *J of Pers and Soc Psych*. 72, 1075–1092.

[17] Kory, J.M. et al. 2013. Robotic learning companions for early language development. *Proc of the 15th ACM on International conference on multimodal interaction* 71–72.

[18] Levenson, R.W. and Gottman, J.M. 1983. Marital interaction: physiological linkage and affective exchange. *Journal of personality and social psychology*. 45, 587–597.

[19] Lutkebohle, I. et al. 2010. The bielefeld anthropomorphic robot head "Flobi." *2010 IEEE International Conference on Robotics and Automation* 3384–3391.

[20] Mizanoor, R. et al. Dynamic Emotion-Based Human-Robot Collaborative Assembly in Manufacturing: The Preliminary Concepts. *hci.cs.wisc.edu*.

[21] Shaer, O. and Hornecker, E. 2010. Tangible user interfaces: past, present, and future directions. *Foundations and Trends in Human-Computer Interaction*. 3, 1-2, 1–137.

[22] Sharma, M. et al. 2013. Communicating affect via flight path Exploring use of the Laban Effort System for designing affective locomotion paths. *Proc of the Int'l Conf on Human-Robot Interaction (HRI), on* 293–300.

[23] Shaver, P. et al. 1987. Emotion knowledge: further exploration of a prototype approach. *Journal of personality and social psychology*. 52, 1061–1086.

[24] Zuckerman, O. 2015. Objects for Change: A Case Study of a Tangible User Interface for Behavior Change. *Ext Abs of the TEI*.