

An Investigation of Responses to Robot-Initiated Touch in a Nursing Context

Tiffany L. Chen · Chih-Hung Aaron King ·
Andrea L. Thomaz · Charles C. Kemp

Accepted: 8 August 2013

© The Author(s) 2013. This article is published with open access at Springerlink.com

Abstract Physical human-robot interaction has the potential to be useful in a number of domains, but this will depend on how people respond to the robot's actions. For some domains, such as healthcare, a robot is likely to initiate physical contact with a person's body. In order to investigate how people respond to this type of interaction, we conducted an experiment with 56 people in which a robotic nurse autonomously touched and wiped each participant's forearm. On average, participants had a favorable response to the first time the robot touched them. However, we found that the perceived intent of the robot significantly influenced people's responses. If people believed that the robot intended to clean their arms, the participants tended to respond more favorably than if they believed the robot intended to comfort them, even though the robot's manipulation behavior was the same. Our results suggest that roboticists should consider this social factor in addition to the mechanics of physical interaction. Surprisingly, we found that participants in our study responded less favorably when given a verbal warning prior to the robot's actions. In addition to these main results, we present post-hoc analyses of participants' galvanic skin responses (GSR), open-ended responses, attitudes towards robots, and responses to a second trial.

Keywords Healthcare · Assistive · Nursing · Touch · Haptics · Intention · Design · User study

1 Introduction

Humans initiate contact with one another to achieve a variety of goals, such as facilitating communication and provid-

ing physical assistance. Robots have the potential to achieve similar goals by initiating physical contact with people, but this type of interaction is fraught with both physical and psychological implications. For example, human skin is an especially important channel for social communication [29], and robot-initiated contact implies that the robot will enter into the person's intimate space [15].

While substantial research has studied how robots can safely operate around people and handle unintended collisions [14], little is known about how a person will respond when a robot intentionally makes contact with the person's body. This type of interaction is especially relevant to healthcare, since caregiving frequently requires that a caregiver initiate contact with the body of a care receiver who is awake and aware. For example, studies of nurse-patient interactions have observed that nurses frequently initiate contact with patients, both to perform tasks that require contact, such as cleaning a person's skin, and to communicate with patients, such as when providing emotional support [7].

2 Overview of Experiment and Main Results

So as to better understand how people respond to robot-initiated touch, we designed and conducted a 2×2 between-subjects experiment with 56 people (14 people per condition) in which a robotic nurse autonomously reached out, touched the participant's arm, moved across the arm, and then retracted. Depending on the condition, the robot verbally indicated before the physical interaction (*warning*) or after (*no warning*) that it intended to clean the participant's arm (*instrumental touch*) or provide comfort (*affectionive touch*). In order to assess participants' responses to these conditions, we took galvanic skin response (GSR) measurements throughout the experiment, administered post-task

T.L. Chen (✉) · C.-H.A. King · A.L. Thomaz · C.C. Kemp
828 W. Peachtree St NW, Suite 204, Atlanta, GA 30308, USA
e-mail: tiffany.chen@gatech.edu

questionnaires, and recorded responses to open-ended questions.

We designed the experiment to test the following two hypotheses:

- *Hypothesis 1*: Participants will find robot-initiated touch more favorable when it is perceived to be instrumental versus affective.
- *Hypothesis 2*: Participants will find robot-initiated touch more favorable when given a verbal warning prior to contact versus no verbal warning.

In agreement with our first hypothesis, we found that participants responded more favorably to the *instrumental touch* than to the *affective touch* conditions. In particular, more people agreed with the statement, “I would have preferred that the robot did not touch my arm.” in the *affective touch* conditions. Nonetheless, all participants let the robot touch them again in a repeated trial. Since the physical behavior of the robot was the same for all trials, our results demonstrate that the perceived intent of robot-initiated touch can significantly influence a person’s subjective response. As such, our results suggest that roboticists should consider this factor in addition to the mechanics of physical interaction.

In contradiction to our second hypothesis, we found that participants tended to respond more favorably to *no warning* than to *warning* conditions. Results from our post-hoc analyses suggest that participants may have become startled by the robot’s voice during the verbal warning, and that the robot’s reach towards the participant’s arm may have served as a warning gesture. However, the underlying reasons for this result remain unclear.

We also report the results of a variety of post-hoc analyses that lend additional insight into the participants’ responses to robot-initiated touch.

3 Related Work

This paper builds upon our initial work communicated via a conference paper [8]. In this article, we provide additional results and analyses, including post-hoc analyses of participants’ galvanic skin responses (GSR), attitudes towards robots, open-ended responses, and responses to a second trial. We also more thoroughly discuss related work, including research published after the submission of our conference paper.

3.1 Nurse-Patient Interaction

Nurse-patient interaction serves as an important source of inspiration for our experiment. It both serves as a motivating application for robots that initiate touch, and a well-studied example of the role of touch in human-human interaction.

Caris-Verhallen et al. observed two types of touch between nurses and patients that they defined as follows: *instrumental touch*, which is “deliberate physical contact” that is necessary in performing a task such as wound dressing; and *affective touch*, which is “relatively spontaneous” and “not necessary for the completion of a task” [7]. In an accompanying study of 165 nurse-patient interactions, researchers observed affective touch in 42 % of the interactions and instrumental touch in 78 % of the interactions [7]. McCann and McKenna report on observations of touching interactions between nurses and older adults in hospice [27]. Most of the observed nurse-initiated touches were on the extremities (arm, hand, leg, foot), and most touches (95.3 %) were instrumental. Touches from nurses on the face, leg, and shoulders were perceived as uncomfortable by patients. Only instrumental touches on the shoulder and arm by a nurse were viewed as comfortable. The authors suggest that misinterpretation of a nurse’s intention may have contributed to patient discomfort during some touches.

In our experiment, we make the same distinction between instrumental and affective touch. By using a robot, we have the distinct advantage of being able to control the physical interaction, and thereby investigate the role of perceived intent through a controlled-laboratory experiment.

3.2 Human-Robot Touch

We classify the initiation of haptic interaction between a human and a robot into three categories: robot-initiated touch, human-initiated touch, and cooperatively-initiated touch. We define robot-initiated touch as a haptic interaction that the robot initiates by making physical contact with the human. Similarly, we define human-initiated touch as a haptic interaction that the human initiates by making physical contact with the robot. We define cooperatively-initiated touch as being a haptic interaction for which the initiator of the touch is ambiguous. For this paper, we also assume that the initiator of contact plays an active role during the interaction episode, while the other entity plays a primarily passive role.

Shaking hands [38] is an example of cooperatively-initiated touch, since both the human and robot can actively move towards each other. When people pet robots, such as Paro [17] or the Haptic Creature robot [44], it is an example of human-initiated touch, since the person actively moves towards a robot and makes physical contact with the robot’s body. Within this paper we focus on robot-initiated touch. Various robotic systems for healthcare involve robot-initiated touch, including facial massage [20], skin care [41], patient transfer [30], surgery [18], and hygiene [19].

There has been some prior work on studying people’s responses to robot-initiated touch. For example, Bickmore et al. have studied users’ perceptions of and responses to affective touch performed by a virtual agent [3]. The virtual

agent included a robotic component capable of pneumatically applying pressure to the user's hand. The user placed his or her hand in the robotic device and held it there. The pressure was initiated by the virtual agent to help convey empathy and comfort. They found marginal trends that suggested that touch increased participants' perceptions of having a working relationship with the agent, if the participants were receptive to touch by humans. The opposite trend was observed for participants who were not receptive to touch by humans.

There has also been a video study that looked at the effect of human-robot touch and robot proactiveness on people's perceptions of a small humanoid robot's "machine-likeness" and dependability [9]. Participants watched videos of a robot and a person interacting. Among other results, participants perceived the robot in the video to be less machine-like when it touched the person while offering to help the person.

Contemporaneous research reported by Nakagawa et al. in [31] is especially relevant to our study. They investigated how a robot making contact with and wiping a participant's hand affects the participant's motivation in a dull task. They compared this interaction with no touch, and the human touching the robot's hand. They found that participants performed the task significantly longer and with more activity when the robot touched and wiped their hands. Furthermore, participants felt that the robot was more friendly when touching them compared with no touch. They plan to use this interaction in healthcare applications, such as encouraging patients during rehabilitation. A number of factors may be responsible for the differences between their results and ours. We discuss these differences in the Discussion and Conclusions section (Sect. 8).

We are unaware of previous research that has directly investigated how the perceived intent of a robot influences a human's subjective response to robot-initiated touch. Likewise, there has been little work on determining cues that robots can use to improve subjective responses to robot-initiated touch, such as a verbal warning.

3.3 Robot Intention

Robot intention plays a critical role in our study. Within this section, we discuss several aspects of robot intention along with related work.

3.3.1 *Must Intentions be Attributed to a Robot?*

With our experiment, we investigated how the perceived intent of a robot influences a person's response to robot-initiated touch. If no intent were attributed to the robot's actions, this inquiry would be inappropriate. Simple machines regularly make contact with people, such as restrain-

ing bars for amusement park rides, automated blood pressure cuffs, and car airbags. Likewise, more complicated machines, such as commercially available massage chairs, autonomously make patterns of contact with peoples' bodies. It seems likely that perceived intent usually does not play a role in these human-machine interactions. Rather, people perceive the devices as mindless mechanisms performing predefined actions.

One could potentially design healthcare robots to reduce the likelihood that people will attribute intentions to them. However, this may not be practical or even possible as healthcare robots become more versatile, perceptive, mobile, dexterous, and communicative in the course of their duties. People tend to attribute intentions, motivations, and emotions to agents that are viewed as being anthropomorphic. The degree to which an agent is viewed as possessing human-like qualities can affect how one predicts what the agent will do in the future as well as what its behaviors mean in the present [11].

Many characteristics can influence a person's tendency to anthropomorphize a machine and attribute intentions to it. People tend to anthropomorphize non-human agents when the agents seem similar to themselves, such as through motion or morphology [11]. For instance, Premack and Premack [33] demonstrated that people will anthropomorphize simple animated 2-dimensional shapes and attribute intentions to them if they move appropriately. Furthermore, mechanical devices such as robots are more readily anthropomorphized when they possess human-like faces and bodies [11, 16].

3.3.2 *The Intentional Stance*

Dennett posits that humans adopt the intentional stance when predicting the future behavior of other systems such as other people [10]. That is, by assuming another person is a rational agent that has beliefs about the world (e.g. there is milk in the refrigerator) and desires or goals of his or her own (e.g. he or she wants to drink milk), one can predict what that person intends to do (e.g. open the refrigerator door, pour a glass of milk, drink the milk, etc.) [1, 10].

Maselli and Altrocchi have stated that "perceived intent is an important determinant of response to another person, and thus attribution of intent is pivotal in understanding interpersonal behavior" [26]. Similarly, the perceived intent of a robot during human-robot interaction (HRI) may influence a person's response to the robot's actions.

3.3.3 *Inferring Intent*

Baird and Baldwin [1] emphasize the roles of high-level knowledge and low-level actions on perceptions of intent. For example, high-level knowledge that a person intends to

tidy up a kitchen that has a sink full of dirty dishes could help an observer recognize that the person intends to clean the dishes. Likewise, observing low-level actions, such as a person directing his or her gaze towards dirty dishes, reaching for a sponge, and reaching for soap, could be used to infer the actor's intent to clean the dishes.

Within our experiment, the experimenters established a high-level context for the scenario by telling participants that the robot was a robotic nurse that would perform various nursing tasks. In conjunction with this context, the robot's speech served to communicate a high-level intention for the robot's actions of either comforting the participant or cleaning the participant's arm. Simultaneously, the robot's motions provided low-level intentions to the participants, such as reaching a hand out and touching the participant's body. We designed our experiment to study the effects of altering the perceived high-level intent of the robot, while keeping the low-level motion the same for all conditions.

3.3.4 Prior Research Involving Robot Intention

Researchers have investigated robot intentions in other contexts. Cakmak et al. considered how a robot can better communicate its intentions in order to improve object handoffs between robots and people [6]. In their study, the experimenter told participants that the high-level intention of the robot was to hand them an object while the experimenter varied the low-level spatiotemporal characteristics of the robot's motion. In our study, we vary the robot's high-level intentions while keeping its low-level actions constant. Wagner and Arkin enabled mobile robots to deceive other robots by giving false signals (heat signatures, sounds, and visual tracks) as to their locations in a game of hide-and-seek. The robots that were hiding communicated a false intent to the robots that were trying to seek them [42].

4 Implementation

In this section, we describe the robot we used in our experiment and the algorithm it used to make physical contact with the participants' arms.

4.1 The Robot

The robot Cody, shown in Fig. 1,¹ is a statically stable mobile manipulator. The components of the robot follow: two arms from MEKA Robotics (MEKA A1), a Segway omnidirectional base (RMP 50 Omni), and a 1 degree-of-freedom



Fig. 1 The robot Cody touches a participant during our experiment

(DoF) Festo linear actuator. Each of the two arms is anthropomorphic with 7 DoF. Each arm joint has a series elastic actuator (SEA) that enables low-stiffness actuation. The robot's wrists are equipped with 6-axis force/torque sensors (ATI Mini40). For this study, we used a custom 3D-printed, spatula-like end effector (7.8 cm × 12.5 cm) which roughly resembles an extended human hand [19]. We cut a towel to fit the shape of the end effector and attached it to the bottom of the end effector. In our experiments, this towel makes contact with the participants' forearms. The towel's material can be interpreted as a material used for cleaning or a compliant exterior for the robot's end effector.

4.2 Touching Behavior Implementation

For implementation details of the touching behavior, please refer to our previous work [19]. During all of the robot's arm motions, the robot's joints were commanded to have low stiffness. For example, when in contact with a participant's forearm, the stiffness of the robot's end effector in the direction normal to the surface of the forearm was less than 60 N/m. For all of the robot's motions in the experiment, the commanded stiffness for the shoulder flexion/extension, shoulder abduction/adduction, shoulder internal/external rotation, elbow flexion/extension, and forearm pronation/supination motions were 20, 50, 15, 25, and 2.5 Nm/rad, respectively. We used position control for the abduction/adduction and flexion/extension motions at the robot's wrist. Even during position control, the wrist joints have significant compliance due to the passive compliance of the SEA springs and cables that connect the SEAs to the joints. For this paper, we attempted to make the touching behavior consistent with both cleaning a person's forearm and providing comfort, so that there would be ambiguity about the purpose of the behavior.

When the robot is in its standby position, its arms and end effectors are pointing down towards the floor. The touching behavior begins by executing what we refer to as the "Init" action. During this action, the robot uses a preprogrammed

¹We obtained IRB approval and participant permission for all photos in this paper.

joint trajectory that moves the left arm to a position where the end effector is 15.4 cm above the mattress surface and directly above the participant's forearm. The robot then moves its end effector downward until the force sensor on the wrist measures a force magnitude ≥ 2 N, indicating that the end effector has made contact with the arm. We designed the arm trajectory so that the "Init" action completed within approximately 7 seconds when tested on a lab member's arm. During the experiment, we recorded the time it took for the robot to perform the "Init" action. The overall mean time for the robot to complete the action across all participants was 6.91 seconds ($SD = 0.10$ sec).

After making contact, the robot performs what we refer to as the "Along" action. During this action, the arm moves the Cartesian equilibrium point (CEP) of the end effector at approximately 4 cm/s. We designed the CEP to travel 14 cm to the left, and then 14 cm to the right along the participant's arm. A bang-bang controller attempts to keep the force magnitude measured by the force sensor on the wrist between 1 and 3 N by moving the CEP down towards the arm or up away from the arm. As a safety precaution, the robot terminates the touching behavior if the measured force magnitude exceeds 30 N. During the "Along" action, the robot exerted an overall mean force magnitude of 2.44 N ($SD = 0.18$ N) across all participants. We also designed this trajectory to be completed in approximately 7 seconds. The overall mean time of the "Along" action across all participants was 6.92 seconds ($SD = 0.02$ sec), and the mean distance the end effector traveled to the left and right was 13.71 cm ($SD = 0.07$ cm) and 13.61 cm ($SD = 0.02$ cm), respectively.

To complete the touching behavior, the robot performs what we refer to as the "Away" action. During this action, the robot lifts its end effector upward, so that it moves away from the person's forearm. The robot then moves its arm back to the standby position. We designed this action to take approximately 7 seconds. The overall mean time for the robot to complete this action across all participants was 6.83 seconds ($SD = 0.08$ sec).

4.3 Safety

Ensuring the safety of a person while interacting with a robot is important during any HRI scenario. Studies in which a robot makes physical contact with a human require special care. We took several precautions when designing the robot's behavior and conducting the study to reduce the chance of injury. First, during the study an experimenter was always prepared to operate a run-stop button if undesirable contact with the robot were observed or anticipated. Second, the robot's arm operated with low joint stiffness and low joint velocities. Third, the robot attempted to keep the magnitude of the force against the participant's arm lower than 3 N.

For comparison, Tsumaki et al. reported that people experienced no pain when a skin care robot applied a downward force of 10 N [41]. Other researchers used a force magnitude threshold of 39.2 N with an oral rehabilitation robot [20]. Various factors could influence the force range that a person would find comfortable, including the contact surface over which the applied force is distributed, and the part of the person's body with which contact has been made. As such, these values provide a coarse comparison with other research.

During the debriefing after the experiment, participants generally reported that the force the robot applied was comfortable. No participants indicated any pain or discomfort during the interaction. Furthermore, in their open-ended responses, several participants reported that the touch was surprisingly light and gentle.

In accordance with the Georgia Institute of Technology Central Institutional Review Board (IRB), we read from a script in order to inform each participant of risks associated with the study, including the potential for undesirable contact with the robot. We also notified each participant that an experimenter would be prepared to use a run-stop button to stop the robot in the event of undesirable contact.

4.4 Galvanic Skin Response (GSR)

In order to provide an objective measure of the participants' arousal during their interactions with the robot, we measured their galvanic skin responses (GSR) using an S220 GSR Sensor from Qubit Systems in Kingston, ON, CA. Several researchers have employed GSR to characterize people's responses during HRI or to enable a robot to respond to a human's affective state [21, 28, 35, 39, 45].

When a person reacts to a stressful situation, the sympathetic nervous system is activated. This activation causes the sweat glands in the palms of the hands and soles of the feet to enlarge, which causes the skin to become more conductive. GSR is linearly correlated with arousal [23] and is generally associated with emotional response [4]. An increase in the voltage reading from the GSR sensor is associated with increased arousal. To use our off-the-shelf GSR sensor, we placed the participant's middle finger and index finger from his or her left hand on the sensor's two electrode plates and secured them with velcro. We attached leads to the electrodes with alligator clips and recorded the voltage reading using the proprietary software Logger Lite. We also recorded timestamps from a clock synchronized with the robot's actions.

We analyzed the GSR signal during the 28-second interval between the baseline recordings described in Sect. 5.1 and shown in Fig. 3. During this time interval, we normalized the GSR signal for each participant to have a value

		Warning Type	
		Warning	No Warning
Touch Type	Instrumental	7 men, 7 women	7 men, 7 women
	Affective	7 men, 7 women	7 men, 7 women

Fig. 2 Experimental design

Warning	Baseline	Speak	Initial	Along	Away	Baseline						
No Warning	Baseline	Initial	Along	Away	Speak	Baseline						
	2 min		7 sec		7 sec		7 sec		7 sec		2 min	

Fig. 3 Timing for *warning* vs. *no warning*

between 0 and 1 inclusive using the following equation, as in [25]:

$$GSR_{norm}(t) = \frac{GSR(t) - GSR_{min}}{GSR_{max} - GSR_{min}}$$

5 Methodology

5.1 Experimental Design

We conducted a gender-balanced, 2×2 between-subjects experiment (see Fig. 2). To test our hypotheses, we defined two independent variables: (1) the type of touch the robot executed (*instrumental* vs. *affective*) and (2) the warning condition (*warning* vs. *no warning*).

In each of the four treatment conditions, the robot executed the same touching behavior described in Sect. 4.2. The only difference between the *instrumental* and *affective* treatment conditions was what the robot said to the participant. The robot used the following utterances:

- *Instrumental, Warning utterance*: “I am going to rub your arm. I am going to clean you. The doctor will be with you shortly.”
- *Instrumental, No Warning utterance*: “I have rubbed your arm. I have cleaned you. The doctor will be with you shortly.”
- *Affective, Warning utterance*: “Everything will be all right, you are doing well. The doctor will be with you shortly.”
- *Affective, No Warning utterance*: “Everything will be all right, you are lying in bed. The doctor will be with you shortly.”

With this design, each participant experienced very similar physical interactions, but associated different intentions with the interaction, depending upon what the robot said. As we describe in detail in Sect. 5.3.4, we asked questions in order to exclude participants who did not interpret the robot’s intentions correctly, which resulted in the exclusion of six

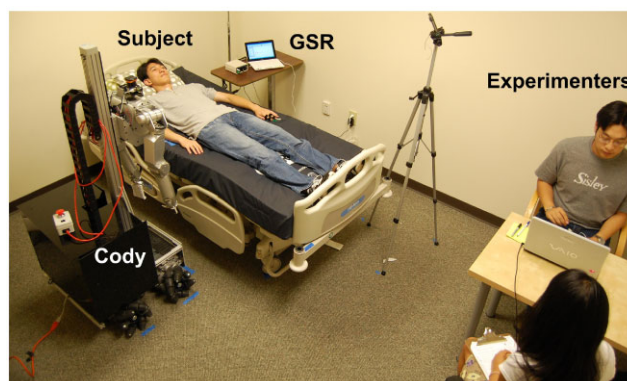


Fig. 4 Experimental setup with a lab member in the patient bed. The two experimenters are shown seated in the bottom-right corner of the image

people. We also controlled the length of time the robot spoke to be approximately 7 seconds for all verbal utterances.

For the *warning* and *no warning* treatment conditions, we changed the timing of when the robot spoke. Figure 3 illustrates the ordering and timing of the robot’s action and speech in the *warning* and *no warning* conditions. For *warning*, the robot spoke before it touched the participant’s arm. For *no warning*, the robot touched the participant’s arm and spoke after the haptic interaction was over (i.e. once it was no longer in contact with the participant’s body). We changed the grammatical construction of the utterances to be appropriate for these two cases.

5.2 Procedure

We recruited 63 students from the Georgia Tech campus through various student email lists, flyers, and word of mouth. We required participants to be at least 18 years of age, a United States citizen, and a native English speaker. We excluded six participants because they did not correctly interpret the robot’s intentions (see Sect. 5.3.4) and one participant due to a software malfunction while collecting her questionnaire data. We assigned participants to each of the four treatment groups on a rolling basis according to gender.

In total, we included the data from 56 of the participants (28 males and 28 females) in the analysis for this paper, ranging in age from 18–29 years ($M = 22.7$, $SD = 2.7$). The self-reported ethnicities of these participants were White (31), Asian (19), African American (2), Hispanic (2), Native Amer./Pac. Islander (1), and Other (1). 87.5 % of the participants were engineering students.

We performed our experiment in the Healthcare Robotics Lab in a $4.3 \text{ m} \times 3.7 \text{ m}$, climate-controlled simulated hospital room (see Fig. 4). We placed a fully functional Hill-Rom 1000 patient bed, an I.V. pole, an overbed table, a living room chair, and a side table in the room. Participants

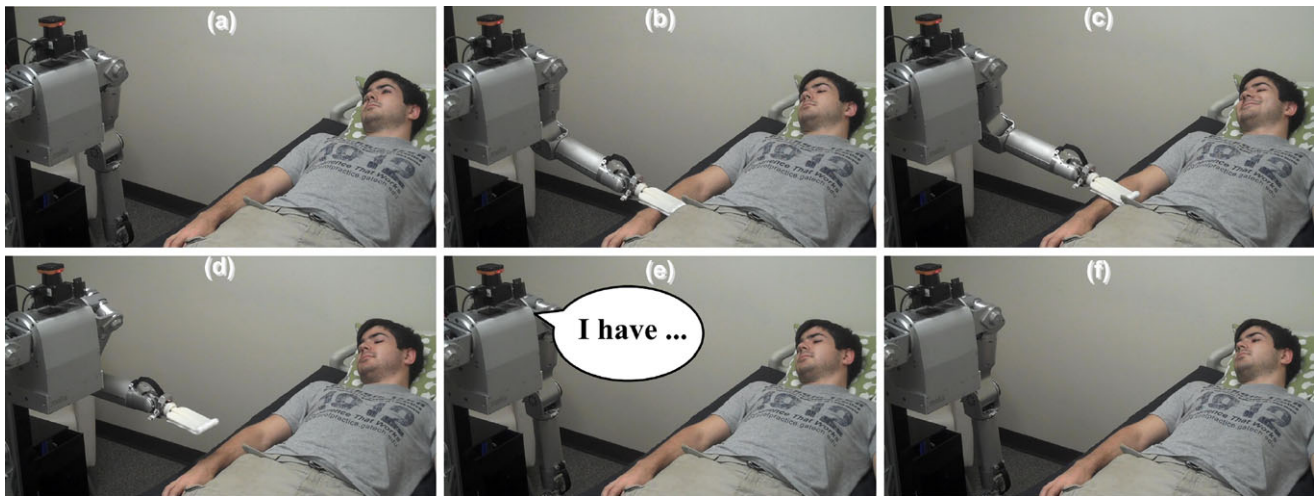


Fig. 5 Cody touches a participant in the *instrumental, no warning* treatment. (a) Baseline. (b) *Init* contact. (c) *Moving Along* the participant's arm. (d) *Lifting Away* from the participant. (e) *Speaking* to the participant (f) Baseline

filled out all paperwork and questionnaires within the simulated hospital room. We placed the robot 17 cm away from the edge of the patient bed.

Two experimenters (the first and second authors of this paper) conducted all of the trials and remained in the room throughout the experiment to ensure the participant's safety. One experimenter, the first author, ran each experiment by reading from a script. While each trial was taking place, the experimenters sat at the far side of the room and looked at a computer monitor and at the robot, rather than at the participant (see Fig. 4). We used a script and the same two experimenters for all trials in order to maintain consistency and avoid confounding factors.

When a participant arrived at the lab, the experimenters welcomed the participant and introduced themselves. Then, the participant signed a consent form, filled out a demographic survey, and filled out a pre-task questionnaire. Afterward, the experimenter explained that the robot was capable of performing several different simulated nursing duties, and that the robot would mimic doing so by gesturing with its arms and end effectors. It is important to note that the participants were unaware that the robot would reach out and make contact with them. Then, the experimenter asked the participant to lay down on the patient bed, and if a female participant was wearing a skirt, the experimenter offered her a blanket to cover her legs. The experimenter then asked the participant to place his or her right arm between two lines of tape marked on the mattress and to place his or her elbow directly on top of a third line of tape on the mattress. This arm placement ensured that the robot would make contact with the person's forearm. If the participant were wearing a long-sleeve shirt or sweater, the experimenter asked the participant to roll up his or her sleeve past the elbow or to remove the sweater, if possible. We asked the participant to

place his or her left arm on the mattress and affixed a galvanic skin response (GSR) sensor to his or her fingers. We collected one minute of baseline data from this sensor and then asked the participant to fill out a brief questionnaire while laying on the bed (measures are detailed in Sect. 5.3).

We then asked the participant to keep his or her head facing a camera during the experiment. Then, we collected 2 additional minutes of baseline GSR data, initiated the robot interaction (described in Sect. 5.1), and collected another 2 minutes of baseline GSR data. We then asked the participant to get off the bed and fill out the post-task questionnaire for trial 1. Next, we asked the participant to lay down in the bed again and performed a repeated trial of the same interaction he or she had just experienced. Then, we asked the participant to fill out a post-task questionnaire for trial 2.

5.2.1 Posture Selection

Since patients are typically in a reclined posture while a nurse performs a bed bath, we selected this posture for our experiment as shown in Fig. 5. The reclined posture may have affected participants' emotional state during the experiment. Previous psychology research has shown that children in a supine position were more fearful than children who were sitting up [22]. Also, physical body posture, specifically slumped, hunched, and relaxed postures, can have an effect on one's emotional state [34]. In our study, we controlled posture across all conditions by asking all participants to recline in the patient bed.

5.3 Measured Subjective Variables

We measured several subjective variables both before and after the participant interacted with the robot by administering pre-task and post-task questionnaires, respectively. In

Table 1 Main effects of touch type on Likert items unrelated to Hypothesis 1

Likert item	Instrumental touch <i>M (SD)</i>	Affective touch <i>M (SD)</i>	<i>F</i> (1, 52)	<i>p</i>	η_p^2
The robot was easy to understand.	5.7 (1.2)	5.9 (1.2)	0.20	.66	0.004
I was confused as to why the robot was touching my arm.	3.0 (2.0)	3.8 (1.6)	3.25	.08	0.06
I was scared when the robot was touching my arm.	2.0 (1.2)	2.4 (1.4)	1.10	.30	0.02
I felt reassured when the robot was touching my arm.	3.6 (1.4)	3.5 (1.2)	0.26	.61	0.005
The robot cares about me.	3.2 (1.5)	2.9 (1.6)	0.60	.45	0.01
The robot was entertaining.	5.4 (1.6)	6.0 (1.3)	2.10	.16	0.04
When I first saw the robot, I thought it would hurt when it touched me.	1.9 (1.2)	2.5 (1.6)	2.2	.14	0.04
The robot looks very strong.	4.8 (1.6)	4.8 (1.5)	0	1	0
The robot looks friendly.	3.2 (1.5)	2.9 (1.3)	0.85	.36	0.02
Interacting with the robot would be more enjoyable if it looked more human-like.	4.9 (1.7)	5.0 (1.6)	0.03	.87	0.001

this section, we describe the measured subjective variables that we use in this paper.

5.3.1 Emotional State

We measured the emotional state of the participants using the Self-Assessment Manikin (SAM) and the Positive and Negative Affect Schedule (PANAS). SAM comprises three 9-point scales that measure arousal, valence, and dominance (also referred to as level of control) using pictorial representations of these dimensions as described in [5, 24]. The Positive and Negative Affect Schedule (PANAS) comprises two 10-word mood scales, where each word is measured on a 5-point scale [43]. Individually, the two scales measure Negative Affect (NA) and Positive Affect (PA), where the lowest possible individual NA or PA score is 10 and the highest is 50. Both SAM and PANAS have been used extensively in psychology and HRI research to measure emotional state [2, 35, 36, 44].

We adapted the text from [5] and [43] for the SAM and PANAS questionnaires we administered. We administered the SAM questionnaire prefaced with the text, “Use these panels to rate your personal reaction OVERALL after the robot finished interacting with you:”. Similarly, we administered the PANAS questionnaire prefaced with the text, “Indicate to what extent you felt the following way OVERALL after the robot finished interacting with you:”.

5.3.2 Custom Likert Item Questionnaire

In addition to assessing the participants’ emotional response, we asked general questions about their experience using 7-point Likert items where 1 = “Strongly Disagree,”

4 = “Neutral,” and 7 = “Strongly Agree.” We asked the following questions pertaining to our two hypotheses:

- LI1 I was confused as to why the robot was touching my arm.
- LI2 It was enjoyable when the robot was touching my arm.
- LI3 I was scared when the robot was touching my arm.
- LI4 I felt reassured when the robot was touching my arm.
- LI5 It was necessary for the robot to touch my arm.
- LI6 I would let the robot touch me again.
- LI7 I would have preferred that the robot did not touch my arm.

The questionnaire included additional questions unrelated to these hypotheses. For completeness, these questions and statistics of the responses to them can be found in Tables 1 and 2.

5.3.3 Negative Attitude Towards Robots Scale (NARS)

We also administered the “Negative Attitude towards Robots Scale” (NARS) which comprises three subscales: S1 which measures negative attitudes towards interactions with robots, S2 which measures negative attitudes towards the “social influence of robots,” and S3 (an inverse scale) which measures positive attitudes towards emotions with robots [32]. NARS has been used to help explain differences found in other measures [40]. We administered NARS during the post-task questionnaire for trial 1 in order to avoid biasing participants prior to their interactions with the robot.

We used methods from [32] to perform our analysis using NARS. We divided the participants into subgroups according to the medians of each of the three NARS subscales S1, S2, and S3. If a participant had an S1 subscale score below

Table 2 Main effects of warning type on Likert items unrelated to Hypothesis 2

Likert item	Warning <i>M (SD)</i>	No warning <i>M (SD)</i>	<i>F</i> (1, 52)	<i>p</i>	η_p^2
The robot was easy to understand.	5.7 (1.2)	5.9 (0.9)	0.45	.51	0.009
It was necessary for the robot to touch my arm	3.9 (2.1)	3.8 (1.6)	0.03	.87	0.001
The robot cares about me.	2.8 (1.5)	3.3 (1.6)	1.22	.27	0.02
The robot was entertaining.	5.7 (1.5)	5.7 (1.5)	0	1	0
When I first saw the robot, I thought it would hurt when it touched me.	2.3 (1.6)	2.1 (1.3)	0.31	.58	0.006
The robot looks very strong.	5.1 (1.4)	4.4 (1.6)	2.55	.12	0.05
The robot looks friendly.	3.0 (1.5)	3.0 (1.3)	0	1	0
Interacting with the robot would be more enjoyable if it looked more human-like.	5.0 (1.6)	4.8 (1.6)	0.24	.63	0.005

Table 3 NARS subscale groupings according to S1, S2, S3 subscale scores. S1 = Negative attitudes towards “situations of interaction with robots,” S2 = Negative attitudes towards the “social influence of robots,” and S3 = Positive attitudes towards emotions with robots

Subscale	Median	Possible range	Subgroup	<i>M (SD)</i>	<i>t</i>	<i>p</i>	Cohen's <i>d</i>
S1	10.5	6–30	Low (<i>n</i> = 28)	8.3 (1.3)	−10.48	< .001	2.89
			High (<i>n</i> = 28)	14.5 (2.8)			
S2	14	5–25	Low (<i>n</i> = 27)	10.7 (2.1)	−11.11	< .001	3.03
			High (<i>n</i> = 29)	16.8 (2.0)			
S3	9	3–15	Low (<i>n</i> = 26)	6.4 (1.4)	−10.3	< .001	2.86
			High (<i>n</i> = 30)	10.8 (1.7)			

the median S1 score, then that participant was placed in the “S1-Low” group. If a participant had an S1 subscale score above the S1 median score, then that participant was placed in the “S1-High” group. We repeated the same process to create the “S2-Low,” “S2-High,” “S3-Low,” and “S3-High” subgroups. We verified that the high and low NARS subscale groupings produced significantly different NARS subscale scores ($p < .001$) for each of the subscales. The results of this verification are shown in Table 3. We used these groups as between-subjects factors in a post-hoc analysis discussed in Sect. 6.3.1.

5.3.4 Manipulation Check

We designed the first two questions of the post-task questionnaire for trial 1 to assess whether participants interpreted the robot's intentions correctly. First, we asked the participant to write down what the robot said to determine if the person correctly heard the robot's speech. Second, we asked the participant to write down why the robot was touching his or her forearm to determine if the person correctly understood the robot's stated intention. We excluded participants who did not pass both of these manipulation checks.

5.4 Expected Outcomes

Within this section, we describe the outcomes we would expect if our hypotheses were true.

5.4.1 Hypothesis 1: Instrumental vs. Affective Touch

Overall, we expect participants to have a stronger preference for the robot not to touch them if the touch were affective as opposed to instrumental (LI7). This is based primarily on the nursing findings described in Sect. 3.1. We also expect participants to experience lower arousal, higher valence, and higher dominance when the robot performs an instrumental touch compared with when it performs an affective touch. Additionally, we expect participants to have higher feelings of positive affect and lower feelings of negative affect when the touch is instrumental. We expect that they would enjoy the touching interaction more (LI2), feel that the touch is more necessary (LI5), and would be more willing to let the robot touch them again when the touch is instrumental (LI6). These expected outcomes correspond with 9 dependent measures.

5.4.2 Hypothesis 2: Warning vs. No Warning

We expect participants to experience lower arousal, higher valence, and higher dominance when they receive a warning

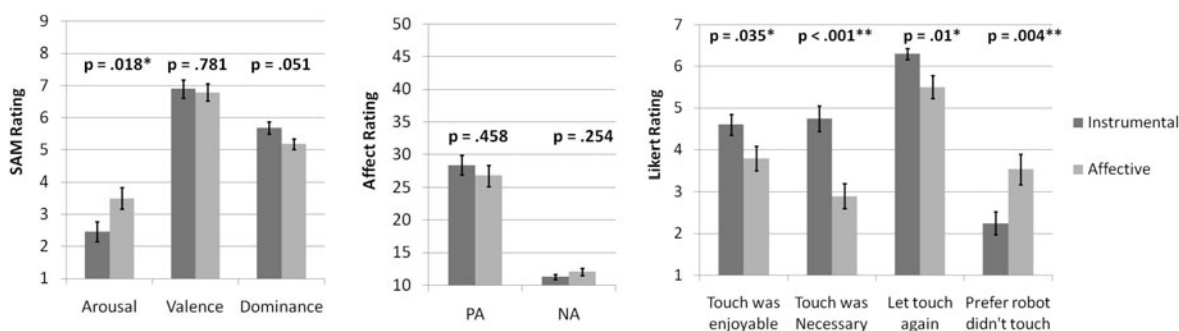


Fig. 6 Main Effects of Touch Type: Participants' subjective responses according to SAM (*left*), PANAS (*middle*), and 7-point Likert items (*right*). (** $p < .0055$, * $p < .05$, Mean and standard error bars shown)

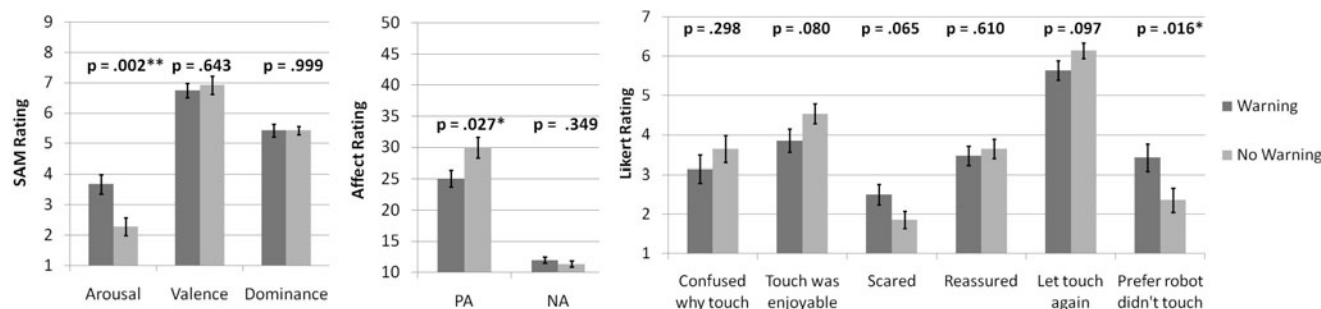


Fig. 7 Main Effects of Warning Type: Participants' subjective responses according to SAM (*left*), PANAS (*middle*), and 7-point Likert items (*right*). (** $p < .0045$, * $p < .05$, Mean and standard error bars shown)

from the robot before it touches them, compared with when the robot touches them before speaking. We also expect participants to have higher feelings of positive affect and lower feelings of negative affect when they receive a warning. We expect participants to enjoy the interaction more (LI2), to be less scared (LI3), to feel more reassured (LI4), and to be more willing to let the robot touch them again (LI6) with a warning. We also expect that with a warning participants would be less confused as to why the robot was touching their arms (LI1), and would be less inclined to prefer that the robot had not touched them (LI7). These expected outcomes correspond with 11 dependent measures.

6 Results

We conducted a two-way, between-subjects analysis of variance (ANOVA) on the subjective data from trial 1 related to the two main hypotheses, and found no significant interactions between the independent variables of *touch type* and *warning type*. Thus, we only discuss the main effects of the independent variables.

Figure 6 shows the main effects of *touch type* on the 9 dependent measures relevant to Hypothesis 1. We denote dependent measures that were significant with $\alpha = .05$ using a single asterisk, *. We denote dependent measures that were

significant with the more conservative Bonferroni adjusted $\alpha = .0055$ (.05/9) using two asterisks, **. The Bonferroni correction reduces the risk of finding significance by chance due to the multiple dependent measures associated with Hypothesis 1 (i.e., Type I errors—false positives).

Similarly, Fig. 7 shows the main effects of *warning type* on the 11 dependent measures relevant to Hypothesis 2. We denote dependent measures that were significant with $\alpha = .05$ using a single asterisk, *. We denote measures that were significant with the more conservative Bonferroni adjusted $\alpha = .0045$ (.05/11) using two asterisks, **.

For completeness, Tables 1 and 2 show the main effects for all other Likert items from the post-task questionnaire. There were no significant interactions between the independent variables for these responses. Furthermore, none of these measures were significant with $\alpha = .05$.

6.1 Hypothesis 1

With respect to the expected outcomes discussed in Sect. 5.4.1, the results were consistent and in support of Hypothesis 1. All 9 dependent measures changed in the anticipated directions, although the changes associated with four of the dependent measures were not statistically significant.

Two dependent measures were significant with the Bonferroni corrected $\alpha = .0055$. Most importantly, more peo-

ple agreed with the statement, “I would have preferred that the robot did not touch my arm.” with affective touch than with instrumental touch (10 participants vs. 1 participant), and there was a statistically significant difference ($F(1, 52) = 9.01$, $p = .004$, $\eta_p^2 = 0.15$) in the responses to this question. This clearly supports Hypothesis 1. Participants also reported that the instrumental touch was significantly more necessary than the affective touch ($F(1, 52) = 18.29$, $p > .001$, $\eta_p^2 = 0.26$). On average, participants viewed the instrumental touch as slightly necessary with a score of $M = 4.8$, $SD = 1.6$ and viewed the affective touch as slightly unnecessary with a score of $M = 2.9$, $SD = 1.6$.

Three other dependent measures were only significant with $\alpha = .05$. Participants were less aroused during the experiment when the robot performed an instrumental touch compared with when it performed an affective touch ($F(1, 52) = 5.92$, $p = .018$, $\eta_p^2 = 0.10$). They also enjoyed the touch more ($F(1, 52) = 4.68$, $p = .035$, $\eta_p^2 = 0.08$) and would be more willing to let the robot touch them again when the touch was instrumental as opposed to affective ($F(1, 52) = 7.05$, $p = .01$, $\eta_p^2 = 0.12$). These results are also consistent with Hypothesis 1.

On average, participants were generally open to allowing the robot interact with them and touch them again, regardless of the touch type. As shown in Fig. 6, participants reported on average that they would let the robot touch them again for both types of touch. Moreover, all 56 participants allowed the robot to touch them in the second trial.

6.2 Hypothesis 2

Surprisingly, with respect to the expected outcomes discussed in Sect. 5.4.2, the results support the contrary assertion that *no warning* results in more favorable subjective responses. 9 out of the 11 dependent measures relevant to Hypothesis 2 changed in the opposite direction from what we anticipated, although the changes associated with six of these dependent measures were not statistically significant. Only the mean rating of confusion changed in the anticipated direction, since people tended to be more confused in the no warning case, albeit not significantly. The average dominance was identical for the *warning* and *no warning* conditions.

Only one dependent measure was significant with the Bonferroni corrected $\alpha = .0045$. Participants were significantly more aroused when the robot warned them prior to contact ($F(1, 52) = 10.71$, $p = .002$, $\eta_p^2 = 0.17$), which is in contradiction to Hypothesis 2.

Two other dependent measures were only significant with $\alpha = .05$. Participants had a higher positive affect rating when the robot did *not* warn them ($F(1, 52) = 5.19$, $p = .027$, $\eta_p^2 = 0.09$). When the robot warned them, participants had a greater preference for the robot not to touch

them ($F(1, 52) = 6.26$, $p = .016$, $\eta_p^2 = 0.11$). These results are in opposition to Hypothesis 2. The changes associated with the remaining eight dependent measures were not significant.

6.3 Post-hoc Analyses

In this section, we describe post-hoc analyses of participants’ negative attitudes towards robots, PANAS scores, a repeated interaction with the robot, GSR, and open-ended responses about the interaction.

6.3.1 NARS

We investigated how participants’ negative attitudes towards robots may have influenced the six dependent measures that showed a significant difference in the first trial (the dependent measures denoted with * and ** in Figs. 6 and 7).

First, we conducted a two-way analysis of variance (ANOVA) for *touch type* and each of the three NARS subscales (S1, S2, and S3). We found no significant interaction between *touch type* and any of the three NARS subscales. We also conducted a two-way ANOVA for *warning type* and the NARS subscales and found no significant interaction. Since we found no interactions, we collapsed across the *touch type* and *warning type* factors and analyzed the main effects of each NARS subscale separately. For each NARS subscale, we used *t*-tests to determine if the six dependent measures were significantly different for participants with high and low NARS subscale results (e.g., S1-Low versus S1-High).

Out of all the *t*-tests for the three NARS subscales and six dependent measures ($3 \times 6 = 18$ *t*-tests), one of the tests was significant with $\alpha = .05$ (Let touch again) while two other tests were marginally significant with $\alpha = .10$ (Prefer no touch, Positive Affect). Table 4 shows the results of these tests. Each of these differences was due to the S3 subscale, which measures positive attitudes towards emotions with robots. Participants who had more positive attitudes towards emotions when interacting with robots (S3-High) were significantly more willing to let the robot touch them again and had marginally higher positive affect. Similarly, participants with less positive attitudes towards emotions when interacting with robots (S3-Low) marginally preferred that the robot not touch them.

6.3.2 PANAS

We compared the PANAS scores from trial 1 with the norms for college students reported in [43]. We performed independent *t*-tests using the mean, standard deviation, and sample size ($n = 660$) statistics for the college students who

Table 4 *t*-tests for selected dependent measures according to NARS subgroups. (Only results significant at the $\alpha = 0.10$ level are shown). S3 = Positive attitudes towards emotions with robots

Subscale	Dependent measure	Subgroup	<i>M</i> (<i>SD</i>)	<i>t</i>	<i>p</i>	Cohen's <i>d</i>
S3	Let touch again	Low	5.5 (1.4)	−2.54	.015	0.73
		High	6.3 (0.8)			
S3	Prefer no touch	Low	3.4 (1.8)	1.968	.054	0.52
		High	2.5 (1.7)			
S3	Positive Affect	Low	25.4 (7.1)	−1.875	.066	0.51
		High	29.4 (8.8)			

Table 5 Comparison of PA scores with the norm reported in [43]. PA score norm: $M = 29.7$, $SD = 7.9$, $n = 660$

Experimental condition	<i>n</i>	PA <i>M</i> (<i>SD</i>)	<i>t</i>	<i>df</i>	<i>p</i>	Cohen's <i>d</i>
Instrumental touch	28	28.4 (8.0)	0.84	29	.41	0.16
Affective touch	28	26.8 (8.5)	1.77	29	.09	0.37
Warning	28	25.1 (7.0)	3.39	30	.002	0.59
No warning	28	30.0 (8.8)	−0.18	29	.86	0.04

Table 6 Comparison of NA scores with the norm reported in [43]. NA score norm: $M = 14.8$, $SD = 5.4$, $n = 660$

Experimental condition	<i>n</i>	NA <i>M</i> (<i>SD</i>)	<i>t</i>	<i>df</i>	<i>p</i>	Cohen's <i>d</i>
Instrumental touch	28	11.3 (2.0)	8.09	46	<.0001	0.66
Affective touch	28	12.1 (3.0)	4.47	35	<.0001	0.51
Warning	28	12.0 (2.6)	5.24	38	<.0001	0.53
No warning	28	11.4 (2.5)	6.58	39	<.0001	0.64

reported how they felt at the moment when they were filling out the questionnaire. The results of these comparisons across conditions are shown in Tables 5 and 6.

The PA scores for participants in the *warning* condition were significantly lower ($p = .002$) than the norm. This result is in line with the results found for Hypothesis 2, where the participants generally did not favor the *warning* condition. All other PA scores were not significantly different than the norm with $\alpha = .05$. The NA scores for participants in all of the conditions were significantly lower than the norm, all with $p < .0001$. This result suggests that participants felt generally less negative than the norm for all experimental conditions. This result may be due in part to the fact that the participants were laying in a bed since low NA is associated with feelings of calmness and serenity [43].

The PA scale has been found to show a time-of-day effect [43] where PA scores tend to rise during the morning, remain steady during the day, and fall during the evening. While we did not explicitly control for the time of day, trials for the various conditions were distributed fairly evenly in time. For our study, 23 people participated in the morning (before noon), 22 participated in the afternoon (between

noon and 6 pm), and 11 participated in the evening (after 6 pm). Notably, participants in the *warning* and *no warning* conditions, for which we found a significant difference in PA, had almost the same distribution across these three times, only differing by one person in the morning, two people in the afternoon, and one person in the evening.

6.3.3 Repeated Trial

We performed a two-way, repeated measures ANOVA on the dependent measures of PA, NA, Arousal, Valence, and Dominance for trials 1 and 2 of the experiment, since these were the only subjective measures that were collected for both trials. The purpose of this analysis was to determine whether the participants' emotional states would change as a result of a second robot-initiated touch interaction. All participants allowed the robot to wipe their forearms in the repeated trial.

We refer to the factor associated with the trial number as *Trial*. The results showed no significant interaction between the *Trial* and *touch type* or *warning type* factors. Thus, we analyzed the *Trial* factor separately. We show the effect due to the *Trial* factor on each of the relevant dependent measures in Table 7. The results show that participants' valence

Table 7 Main effect of the *Trial* factor on participants' self-reported emotional state

Dependent measure	Possible range	Trial 1 <i>M</i> (<i>SD</i>)	Trial 2 <i>M</i> (<i>SD</i>)	<i>F</i> (1, 52)	<i>p</i>	η_p^2
Arousal	1–9	3.0 (1.8)	2.7 (1.7)	2.74	.10	0.05
Valence	1–9	6.8 (1.4)	6.3 (1.4)	16.11	<.001	0.24
Dominance	1–9	5.4 (1.0)	5.4 (0.9)	0.07	.79	0.001
Positive affect	10–50	27.6 (8.2)	22.9 (7.1)	55.15	<.0001	0.52
Negative affect	10–50	11.7 (2.5)	12.9 (5.4)	3.4	.07	0.06

Table 8 Comparison of PANAS scores from repeated trials with the norms reported in [43]. PA score norm: $M = 29.7$, $SD = 7.9$, $n = 660$; NA score norm: $M = 14.8$, $SD = 5.4$, $n = 660$

Trial and score type	<i>n</i>	Score <i>M</i> (<i>SD</i>)	<i>t</i>	<i>df</i>	<i>p</i>	Cohen's <i>d</i>
Trial 1, PA	56	27.6 (8.2)	1.85	64	.07	0.27
Trial 1, NA	56	11.7 (2.5)	7.85	106	<.0001	0.59
Trial 2, PA	56	22.9 (7.1)	6.82	67	<.0001	0.87
Trial 2, NA	56	12.9 (5.4)	2.53	65	.01	0.35

and PA scores significantly decreased from trial 1 to trial 2 ($p < .001$ and $p < .0001$, respectively), indicating that they became significantly less happy. Similarly, participants' NA scores increased with marginal significance ($p = .07$). Participants also became less aroused from trial 1 to trial 2 with marginal significance ($p = .10$), while their feelings of dominance did not significantly change.

We also compared the PANAS scores for trials 1 and 2 with the norms from [43] across all participants. The results from this analysis are shown in Table 8. The participants' PA scores were not significantly different than the norm after trial 1 ($p = .07$), but were significantly lower than the norm ($p < .0001$) after trial 2. Also, in trial 1, the participants' NA scores were significantly lower than the norm ($p < .0001$), and remained significantly lower than the norm after trial 2 albeit to a lesser degree ($p = .01$) than in trial 1. These results indicate that participants, on average, were less positive than the norm in trial 2 compared with trial 1, but were less negative than the norm in both trials.

6.3.4 GSR

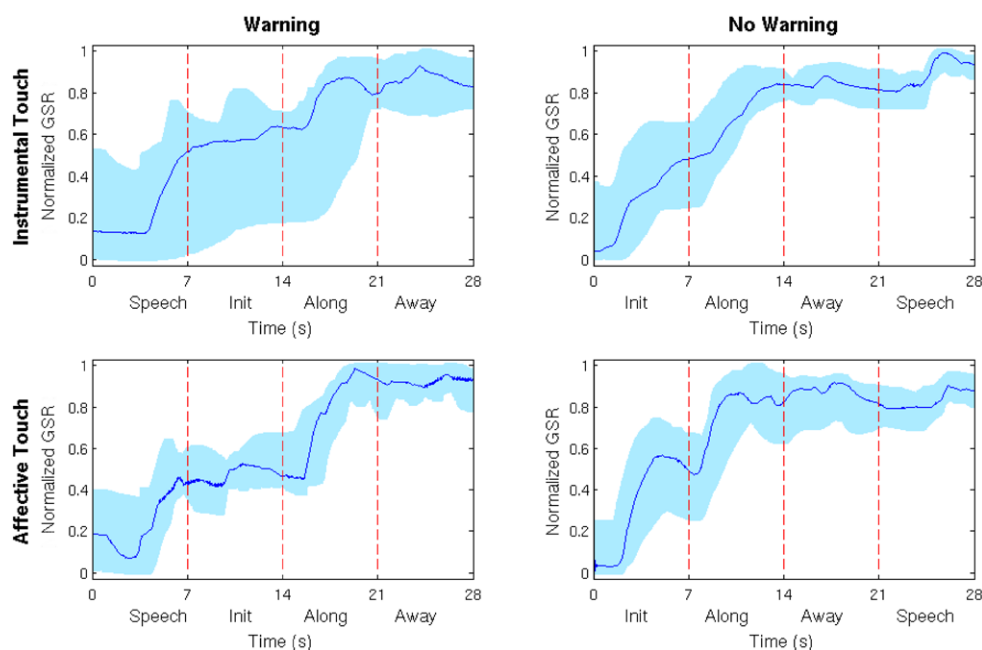
The results from analyzing the participants' GSR suggest that the robot's first action served as a form of warning, whether it was spoken (*Speech*) or gestural (*Init*). After the participants had been warned via speech or gesture, these non-contact actions did not result in further ascents in arousal. In contrast, the actions associated with the *Along* interval, which involved contact with the participant's body, resulted in ascents in arousal regardless of the spoken or ges-

**Fig. 8** Various facial expressions during the *Along* time interval

tural warning. For reference, Fig. 8 shows examples of participants' facial expressions during the *Along* interval.

Figure 9 shows the median of the normalized GSR across the participants for their first interaction with the robot according to treatment group. We omitted the GSR data from 10 participants due to errors in the recorded timestamps, from 5 participants due to erroneous measurements (excessive high-frequency content), and from 1 participant due to signal drop out during the interaction. Consequently, Fig. 9 shows GSR readings from a total of 40 participants. Specifically, we used data from 10 participants each in the *Instrumental touch*, *Warning* and *Affective touch*, *Warning* groups, 8 participants in the *Instrumental touch*, *No Warning* group, and 12 participants in the *Affective touch*, *No Warning* group.

Fig. 9 Normalized GSR across each treatment. *Dark blue line* shows median of normalized GSR signal. *Light blue area* shows data contained within the 25th and 75th percentiles. *Red dashed vertical lines* show timing of the *warning* and *no warning* conditions (shown previously in Fig. 3). $n = 10$ for *Instrumental touch, Warning* and *Affective touch, Warning*; $n = 8$ for *Instrumental touch, No Warning*; and $n = 12$ for *Affective touch, No Warning* (Color figure online)



Statistically Significant Ascents The GSR curves in Fig. 9 are of a “type 3” pattern curve where there are not distinct peaks following each stimulus but are instead subsequent “ascents” [4]. This pattern may arise when stimuli are placed close enough such that a descent in the GSR is not produced between stimuli. The latency between the onset of a stimulus and the onset of a GSR response is typically between 1 and 2 seconds [4].

Due to our experimental design, participants began the human-robot interaction part of the experiment represented in Fig. 9 in a relaxed state with a low-level of arousal and low GSR readings. From the GSR trends, we observe that the participants’ arousal tended to increase throughout the interaction.

Table 9 shows the results of pairwise t -tests comparing the normalized GSR values at the end and beginning of each interval for each treatment. Notably, during the *Along* interval the GSR signal had a significant ascent (increase in GSR) in all four conditions. This observation indicates that the instant of contact and motion along the forearm were arousing under all conditions, regardless of the interaction that had already taken place. In contrast, the robot’s retraction of its arm during the *Away* interval did not result in a significant change in GSR in any of the treatment conditions.

A significant ascent was associated with the first interval under all conditions. When the *Init* interval was the first action during the human-robot interaction (for the *No Warning* condition), there was a significant increase in GSR during that interval across both of the *touch types*. This result was not observed for the *Init* interval under the *Warning* condition, where *Init* was not the first action of the robot, but instead occurred after *Speech*. The *Speech* interval corre-

sponded with a significant increase in GSR for all four conditions. However, the ascent was much larger in the *Warning* condition (across both *touch types*), when the robot’s utterance was the first action the robot performed, compared with the ascent in the *No Warning* conditions, where the robot’s *Speech* was the last interval of the interaction.

Intervals with Similar GSR Change We also compared intervals to one another. For our analysis of the *Init* and *Along* intervals, we combined the GSR data from the two *no warning* conditions, since these experimental conditions are the same prior to the *Speech* interval.

We performed three independent samples t -tests to compare the GSR difference for the combined *no warning* condition ($M = 0.32$, $SD = 0.22$), the *instrumental touch, warning* condition ($M = 0.28$, $SD = 0.22$), and the *affective touch, warning* condition ($M = 0.36$, $SD = 0.17$). We found no statistically significant difference in the GSR ascents associated with the *Along* interval under these three conditions. The results for these tests are as follows: *instrumental touch, warning* vs. *affective touch, warning*: $t(18) = -0.97$, $p = .34$, $d = 0.43$; *instrumental touch, warning* vs. *no warning*: $t(28) = -0.54$, $p = 0.60$, $d = 0.19$; and *affective touch, warning* vs. *no warning*: $t(28) = 0.49$, $p = .63$, $d = 0.20$.

The GSR signal had a significant ascent in the first interval for all conditions. The first interval was *Speech* in the *warning* conditions and *Init* in the *no warning* conditions. The change in GSR signal for these two intervals was not significantly different (*Speech, Warning*: $M = 0.26$, $SD = 0.20$, *Init, No Warning*: $M = 0.33$, $SD = 0.20$, $t(38) = -1.2$, $p = .25$, $d = 0.36$). This result suggests that

Table 9 Pairwise *t*-tests of normalized GSR at beginning and ending of each interval for each treatment. The mean and standard deviation of the difference between the normalized GSR value at the end of the interval and the beginning of the interval is shown

Interval	Treatment	GSR difference <i>M</i> (<i>SD</i>)	<i>t</i>	<i>df</i>	<i>p</i>	Cohen's <i>d</i>
<i>Init</i>	<i>Instrumental, Warning</i>	0.04 (0.14)	0.93	9	.38	0.29
	<i>Instrumental, No Warning</i>	0.30 (0.21)	4.07	7	.005	1.43
	<i>Affective, Warning</i>	0.07 (0.15)	1.48	9	.17	0.47
	<i>Affective, No Warning</i>	0.36 (0.20)	6.27	11	<.0001	1.80
<i>Along</i>	<i>Instrumental, Warning</i>	0.28 (0.22)	3.97	9	.003	1.27
	<i>Instrumental, No Warning</i>	0.37 (0.27)	3.91	7	.006	1.37
	<i>Affective, Warning</i>	0.36 (0.17)	6.95	9	<.0001	2.12
	<i>Affective, No Warning</i>	0.30 (0.20)	5.2	11	<.001	1.50
<i>Away</i>	<i>Instrumental, Warning</i>	0.04 (0.09)	1.23	9	.25	0.44
	<i>Instrumental, No Warning</i>	-0.02 (0.09)	-0.56	7	.59	0.22
	<i>Affective, Warning</i>	0.002 (0.08)	0.06	9	.95	0.03
	<i>Affective, No Warning</i>	0.007 (0.14)	0.17	11	.87	0.05
<i>Speech</i>	<i>Instrumental, Warning</i>	0.26 (0.24)	3.37	9	.008	1.08
	<i>Instrumental, No Warning</i>	0.11 (0.11)	2.88	7	.02	1.00
	<i>Affective, Warning</i>	0.26 (0.17)	4.8	9	<.001	1.53
	<i>Affective, No Warning</i>	0.07 (0.07)	3.10	11	.01	1.00

the robot's actions during the first interval resulted in comparable increases in arousal, regardless of whether the action was the robot speaking or the robot moving its arm.

Additionally, neither the *Speech* interval nor the *Init* interval was associated with an ascent when they occurred at other times during the interaction. The increase in GSR was significantly larger when *speech* was the first interval of the interaction as opposed to when it was the last interval (*Speech, Warning*: $M = 0.26$, $SD = 0.20$, *Speech, no warning*: $M = 0.08$, $SD = 0.09$, $t(38) = 3.5$, $p < .002$, $d = 1.19$). Similarly, the change in GSR was significantly larger when the *Init* interval was the first interval as opposed to the second (*Init, No Warning*: $M = 0.33$, $SD = 0.20$, *Init, warning*: $M = 0.05$, $SD = 0.14$, $t(38) = -5.2$, $p < .001$, $d = 1.66$).

Many factors can influence GSR, including posture, age, temperature and lighting [4]. However, these factors were consistent across all trials.

6.3.5 Experience with Robots

26 participants (46 %) responded "Yes" to the question "Do you have any experience with robots?" The examples of robots with which they reported having experience included toys, LEGO Mindstorms, iRobot Roomba, Boe-Bot, the Philips iCat, the Willow Garage PR2, an autonomous ground vehicle, and a bioreactor for tissue engineering. We compared the responses of participants who reported having robot experience ($N = 26$) with the participants who re-

ported having no robot experience ($N = 30$). We first performed independent samples *t*-tests to compare their responses for each dependent measure used to test hypotheses 1 and 2 (12 measures total) and for each of the NARS subscales. No test was significant with $\alpha = .05$. We then analyzed the effect size using Cohen's *d*. [12] recommends $d = .41$ be considered a minimum effect size of practical significance for social science data. Only one measure met this cutoff, which was question LI4 ("reassured") with $d = .42$ where those who reported no experience with robots felt slightly more reassured ($M = 3.8$, $SD = 1.2$) than those who reported previous experience with robots ($M = 3.3$, $SD = 1.3$). This suggests that experience with robots did not have a substantive effect on the results of our study.

6.3.6 Open-Ended Responses

After the first touching interaction with the robot, we asked the participants to answer two open-ended questions in the post-task questionnaire. For our analysis, we read through all of the participants' responses for each of the open-ended questions. Then, we created categories for the types of responses people made. We present our results with respect to these categories. We also provide sample quotations from the responses to give the reader an idea of the types of comments the participants made.

Question #1: What Would You Suggest we Change About the Robot in Order to Make the Interaction More comfortable?

We grouped the responses to question #1 according to comments concerning: (1) the robot's voice, (2) the robot providing a warning prior to touch, (3) the robot's movement, (4) the appearance of the robot, (5) the robot saying more, and (6) the design of the end effector of the robot.

21 out of the 56 participants mentioned that they wanted to change the robot's voice in some way. 13 out of these 21 participants (62 %) were in the *Warning* treatment group. 9 out of these 21 participants wanted the voice to be more human-like, and 3 out of these 21 participants wanted the voice to be friendlier, while 3 other of these 21 participants simply wanted the voice to be "better" or to "improve" it with no other specific description as to how to improve it. No participants reported that the choice of the female voice should be changed.

15 of the 56 participants reported that they would like to have had some sort of warning prior to the robot touching them. 11 of these 15 participants (73 %) were in the *no warning* treatment group. 3 of the 4 participants who were in the *warning* treatment group noted that when the robot warned them, its voice startled them. Specifically, 1 of these 3 participants mentioned that the speech was surprising, since the robot had been silent, while another 1 of these 3 noted that the robot had a "thundering voice" and that the volume of the voice should be lowered. Being startled by the robot's voice may have contributed to the participants' higher arousal ratings and our unexpected results. 1 of the 4 participants in the *warning* group suggested that the robot make an initial small movement before it touched the person.

21 of the 56 participants made suggestions to change some aspect of the robot's appearance. Specifically, 7 of these 21 wanted the robot to have a head or a face; 5 of these 21 simply wanted the robot to look more "friendly"; and 7 of these 21 wanted the robot to be less "metallic" looking or less mechanical. 11 of these 21 participants expressed that they wanted the robot to have more humanoid characteristics and specifically mentioned some form of the word "human."

5 of the 56 participants indicated that they would have liked if the robot had spoken more. 4 of these 5 were in the *no warning* treatment group. 1 of these 4 participants suggested that the robot introduce itself, while 2 other of these 4 participants suggested that the robot should say more about the context of the situation and should give more indication of what was about to happen. 1 of these 4 participants simply suggested that the robot engage in "small talk." The single participant in the *warning* treatment group desired that the robot provide a longer explanation about the cleaning.

4 of the 56 participants made design suggestions for the robot's end effector. 1 of these 4 participants wanted the cloth on the end effector to be replaced by a more "human-like replacement" such as a rubbery material and a warming element. Another 1 of these 4 participants echoed the desire for a warming element if touch was involved. 2 of these

4 participants suggested to make the cloth softer and less rough.

1 of the 56 participants stated that: "I could see a robot performing tasks, not necessarily providing emotional support like comforting someone."

Question #2: What Are Your Overall Impressions of the Experiment? Many of the responses to question #2 were in line with the suggestions discussed already for question #1, and included suggestions to change the robot's voice and appearance, possibly to be more human-like.

1 participant stated that "I think that I would be comfortable having a healthcare robot interact with me in a doctor's office." On the other hand, another participant stated it was "not the same as having a human nurse" and another stated that she was "somewhat doubtful that interacting with the robot would be comforting in the same way as with a human."

6 of the 56 participants noted that they were surprised how lightly the robot touched them and that it was more gentle than they had expected. 1 of these 6 participants stated that the light touch was "reassuring." This participant was in the *Instrumental touch, No Warning* treatment group.

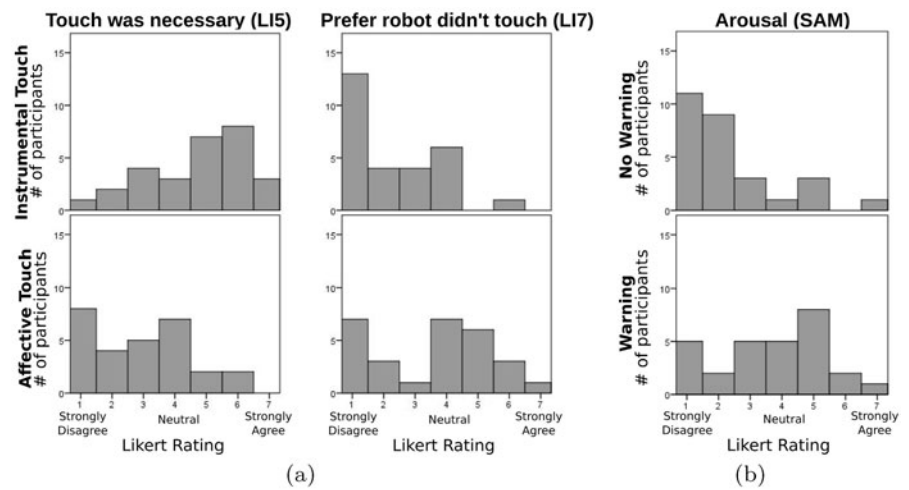
On a similar note, 1 of the 56 participants stated that "I was surprised that I in fact felt more calm after interacting with the robot." This participant was in the *Affective touch, Warning* treatment group.

8 of the 56 participants noted that they expected to do more in the experiment or that the experiment would have more elements of interaction. 6 other of the 56 participants explicitly stated that they wanted more interaction with the robot.

9 of the 56 participants expressed confusion about parts of the experiment. Specifically, 5 of these 9 participants stated that they were confused about questions in the questionnaire. 1 of these 9 participants expressed that the robot's voice was difficult to understand. Although all 56 participants passed the manipulation check, 2 of the 9 participants stated that they were confused about why the robot was touching them. Of these 2 participants, 1 was in the *no warning, instrumental touch* group and 1 was in the *warning, affective touch* group. 1 of the 9 participants wondered whether there would be a second interaction immediately following the first.

5 of the 56 participants expressed negative feelings towards parts of the experiment. 1 of these 5 participants stated that he felt odd that the robot was trying to comfort him. Another of these 5 participants stated that he "felt weird to be laying on the bed," while another of these 5 participants felt uncomfortable during the period of waiting. Another of these 5 participants stated that being told not to move made her worried about doing something wrong during the experiment. The last of these 5 participants stated that the touch

Fig. 10 Histograms of Significant Dependent Measures According to Main Effect of (a) Touch Type, (b) Warning Type



itself felt “weird” but that a warning would have made the touch less awkward. This participant was in the *Instrumental touch, No Warning* treatment group.

7 Limitations

Further research will be required to determine the generality of our results. We carefully controlled factors such as the robot’s appearance, the robot’s motions, the location where contact was made on the person’s body, and the person’s posture. Any one of these or other factors, such as long-term interaction with the robot, the person’s cultural background, or previous experience with robots could potentially have a significant influence on a person’s response. For example, the participants in this study were predominantly engineering college students, which may limit generalization of our results to other populations.

Likewise, the participants were in a simulated scenario. Patients who actually require care or would benefit from comfort might respond differently. For instrumental touch, patients who require care might respond more positively, since the touch would truly be instrumental. For affective touch, it is unclear if people who would benefit from comfort would respond more or less favorably.

Additionally, participants were under informed consent, and, hence, knew they were part of an experiment. As such, the Hawthorne effect may have been a factor in our results. Since we carefully controlled the experiment, we would not expect this to be a confounding factor for our results based on comparisons across conditions, such as our finding that perceived intent can significantly influence a person’s response to robot-initiated contact. On the other hand, participants’ speculations about the nature of the experiment could potentially have influenced them to respond more positively to the interaction. That said, under the affective conditions, a

large number of people (10 out of 28) agreed with the statement, “I would have preferred that the robot did not touch my arm.” (see Fig. 10). This demonstrates that many participants were willing to provide negative responses.

During recruitment, we specifically mentioned that participants would be interacting with a robot, but we did not indicate that the robot would make physical contact with them nor did we state that the robot would act as a nurse. Nonetheless, our recruiting method would be less likely to enroll people who are averse to interacting with robots.

8 Discussion and Conclusions

We have presented results from our study in which a human-scale robot using a compliant arm autonomously made contact with the forearms of 56 human participants without incident or reported discomfort. On average, regardless of the treatment, participants had a favorable experience in the first trial as indicated by measures such as valence, positive affect, and negative affect, as well as Likert items about perceived safety, fear of the robot, and willingness to have the robot touch them again. In general, these results suggest that robot-initiated touch can be a successful form of human-robot interaction in the context of healthcare. More specifically, in this study we investigated how two factors influence the response of participants to robot-initiated touch. We selected these factors based on their relevance to human-human interaction in the context of nursing.

8.1 Perceived Intent

Our study demonstrates that perceived intent can be a significant factor in how people respond to robot-initiated touch. For all trials in our experiment, the robot executed the same touching behavior, which resulted in consistent

physical interaction with the participants. Significant variation in responses resulted from distinct interpretations of the robot's intentions rather than physical differences in the interaction. Specifically, participants responded less favorably when they believed the robot touched them to comfort them (affective touch) versus when they believed the robot touched them to clean their arms (instrumental touch). This was most evident in participants' agreement with the statement, "I would have preferred that the robot did not touch my arm." 10 participants from the affective touch conditions agreed with the statement versus 1 participant from the instrumental touch conditions, and the average response to this question was significantly different for affective touch and instrumental touch $p = .004$).

In our study, the robot touched a relatively innocuous location on the participant's body. We would anticipate a stronger effect size if the robot were touching a more sensitive part of the body, such as during an actual bed bath. We would also expect the role of perceived intent to play a larger role during contact with more sensitive locations. For example, if a nursing robot made contact with a particularly sensitive part of a patient's body, the patient's response would likely depend on whether or not he or she believed the contact was a mistake, was intended to provide comfort, or was intended to achieve a medical goal.

Exploring ways to reinforce desired interpretations of robot-initiated touch could be a worthwhile direction for future research. In our study, we used the robot's speech, the actions of its arm, and the nursing scenario to convey intent. Many other cues, including high-level, low-level, implicit, and explicit cues, could plausibly be used to influence perceived intent. Alternatively, developing robots to which people are unlikely to attribute intentions may be an effective approach. As we discussed in the related work section (Sect. 3.3), however, this may not be feasible as healthcare robots become more advanced and less specialized.

8.2 Instrumental vs. Affective Touch

Studies of interactions between human nurses and human patients have found that patients respond more favorably to instrumental touch. Similarly, with our robot in a simulated healthcare task, participants responded much more favorably to instrumental touch than to affective touch. However, the extent to which this result would generalize to other robots in other scenarios remains an open question.

It may be possible to create a nursing robot from which people would welcome comforting touch and instrumental touch. For example, the robot's appearance and behavior might be altered to better match the tasks [13, 37]. As we described earlier, Nakagawa et al. have recently shown that a robot touching and wiping the hand of a participant can positively motivate the participant in a task [31]. Many factors

may have led to the positive responses they observed. They used a small, cute, child-sized robot designed specifically for social interaction, while we used a human-scale mobile manipulator that we primarily designed to perform instrumental healthcare-related tasks. Their participants were seated upright and looking down at their robot, which was placed on a table. This dominant posture, similar to an adult interacting with a child, is in contrast to the supine posture of our participants who were looking up at the robot from a hospital bed. Cultural differences may also have been a factor, since their participants were from Japan while ours were native English speakers residing in the Atlanta area of the USA. Additionally, their robot asked the participants to first hold its right hand, which resulted in human-initiated touch prior to robot-initiated touch. In our study, the robot made first contact with participants.

All of these factors might make a difference in people's responses to robot-initiated touch, and some of them would be difficult or impossible for a roboticist to control. For example, robotic nurses may need to interact with people of many cultures who are in a supine posture. Likewise, the demands of instrumental tasks will place requirements on the robot's design, and may call for human-scale or larger robots.

Fortunately, with respect to the goal of instrumental touch, our results suggest that favorable responses can be achieved with a robot lacking strong social design elements. Participants from the instrumental touch conditions had a generally positive response to the first trial. Notably, only 1 participant out of 28 agreed with the statement "I would have preferred that the robot did not touch my arm." (see Fig. 10). Whether or not the addition of social design elements would improve responses remains an open question. Some healthcare robots have integrated social design elements, such as the large teddy-bear-like RIBA robot that is designed for lifting patients [30]. One potential risk is that people might not respond well to some socially-oriented design choices in a large healthcare robot. The open-ended responses from our participants indicate that adding human-like characteristics may be beneficial.

8.3 Warning, Warning

We found that participants tended to respond more favorably when no verbal warning was given by the robot prior to contact. The open-ended responses and GSR data suggest that the movement of the robot as it reached out towards the person served as a form of gestural warning, and that it was preferred to the robot's spoken warning. The robot's unexpected physical movement and contact with the person after a long period of stillness may have been less jarring than the robot's unexpected speech after a long period of silence. It seems likely that factors such as the velocity of the

arm movement and the loudness of the speech play a role in this type of interaction. As suggested in participants' open-ended responses, having the robot speak to the person ahead of time in a more natural manner, or otherwise reducing surprise, might lead to different results.

Another interpretation of our results could be that leaving the intent behind robot-initiated touch ambiguous while the interaction is occurring leads to more favorable responses. The robot explicitly stating its high-level intentions may have caused apprehension for the participants. This would seem to go against common bedside manner as practiced by human nurses. The speech prior to the robot's actions may have also resulted in a stronger tendency for participants to anthropomorphize the robot, which may have influenced their responses to the robot. Similarly, no warning may have resulted in participants perceiving the robot as more machine-like. Interestingly, in the open-ended responses, 11 participants who had not been warned stated that they would have liked to have been warned.

Further research will be required to confidently interpret these surprising results. For now, our results suggest that verbal warnings prior to contact should be carefully designed, if used at all, and that gesture can serve as a form of warning. When a robot should perform low-level communication of intention via gestures versus high-level communication of intention via speech remains an open question.

8.4 Negative Attitude Towards Robots Scale (NARS)

Our post-hoc analyses suggest that participants who have less positive attitudes towards emotions in interactions with robots as measured by NARS respond less favorably to robot-initiated touch. These results suggest that robot-initiated touch is related to emotional elements of interaction, and S3 subscale scores could be informative when a robot engages a person in such an interaction. However, we administered the NARS instrument after the experiment in order to avoid biasing the participants. As such, we found an association in the responses to NARS and other measures after the interaction, but we do not know if NARS responses collected prior to the interaction would have been predictive.

8.5 The Repeated Trial

Perhaps the most notable result from the repeated trial is that all participants allowed the robot to touch them again. Participants responded less favorably to the second interaction with the robot. It is likely that the nature of the repeated experiment resulted in people responding less favorably. The total experiment with both trials took approximately 45 minutes, included long periods of waiting, and involved filling out numerous questionnaires. In addition, the second trial was identical to the first, and consequently could have been less interesting to the participants.

8.6 Robot-Initiated Touch for Healthcare and Medicine

Our results provide evidence that robot-initiated touch can be a practical form of human-robot interaction. Procedures associated with health and medicine often entail discomfort, such as when having blood drawn, receiving a bed bath, undergoing dental work, or being in the confined space of an MRI scanner. Our results suggest that robot-initiated touch can play a role in healthcare but that it need not be distressing for patients. However, further research will be required to generalize our results to real patients, including patients from other demographics who have not explicitly chosen to interact with a robot.

Acknowledgements We thank Zhengqin Fan, Advait Jain, Mrinal Rath, Kate Tsui, and all the participants. We thank Dr. William R. Delaune for his advice on statistical analysis. We gratefully acknowledge support from NSF awards IIS-1150157 and EFRI-1137229, and an NSF GRFP.

Open Access This article is distributed under the terms of the Creative Commons Attribution License which permits any use, distribution, and reproduction in any medium, provided the original author(s) and the source are credited.

References

- Baird J, Baldwin D (2001) Making sense of human behavior: action parsing and intentional inference. In: *Intentions and intentionality: foundations of social cognition*, pp 193–206
- Bethel C, Burke J, Murphy R, Salomon K (2007) Psychophysiological experimental design for use in human-robot interaction studies. In: *2007 international symposium on collaborative technologies and systems (CTS)*, pp 99–105
- Bickmore TW, Fernando R, Ring L, Schulman D (2010) Empathic touch by relational agents. *IEEE Trans Affect Comput* 1(1):60–71
- Boucsein W (2011) *Electrodermal activity*. Springer, Berlin
- Bradley M, Lang P (1994) Measuring emotion: the self-assessment manikin and the semantic differential. *J Behav Ther Exp Psychiatry* 25(1):49–59
- Cakmak M, Srinivasa S, Lee M, Kiesler S, Forlizzi J (2011) Using spatial and temporal contrast for fluent robot-human hand-overs. In: *Proceedings of the 6th ACM/IEEE international conference on human-robot interaction*, pp 489–496
- Caris-Verhallen WM, Kerkstra A, Bensing JM (1999) Non-verbal behaviour in nurse-elderly patient communication. *J Adv Nurs* 29(4):808–818
- Chen T, King C, Thomaz A, Kemp C (2011) Touched by a robot: an investigation of subjective responses to robot-initiated touch. In: *Proceedings of the 6th ACM/IEEE international conference on human-robot interaction*, pp 457–464
- Cramer HS, Kemper NA, Amin A, Evers V (2009) The effects of robot touch and proactive behaviour on perceptions of human-robot interactions. In: *Proceedings of the 4th ACM/IEEE international conference on human-robot interaction*, pp 275–276
- Dennett D (2007) True believers: the intentional strategy and why it works. In: Haugeland J (ed) *Mind design II*. MIT Press, Cambridge
- Epley N, Waytz A, Cacioppo J (2007) On seeing human: a three-factor theory of anthropomorphism. *Psychol Rev* 114(4):864

12. Ferguson CJ (2009) An effect size primer: a guide for clinicians and researchers. *Prof Psychol, Res Pract* 40(5):532
13. Goetz J, Kiesler S, Powers A (2003) Matching robot appearance and behavior to tasks to improve human-robot cooperation. In: *Proceedings of the 12th IEEE international workshop on robot and human interactive communication (RO-MAN)*, pp 55–60
14. Haddadin S, Albu-Schäffer A, Hirzinger G (2010) Safety analysis for a human-friendly manipulator. *Int J Soc Robot* 2(3):235–252
15. Hall ET (1966) *The hidden dimension*. Anchor Books, New York
16. Hegel F, Krach S, Kircher T, Wrede B, Sagerer G (2008) Understanding social robots: a user study on anthropomorphism. In: *Proceedings of the 17th IEEE international symposium on robot and human interactive communication (RO-MAN)*, pp 574–579
17. Kidd CD, Taggart W, Turkle S (2006) A sociable robot to encourage social interaction among the elderly. In: *Proceedings 2006 IEEE international conference on robotics and automation (ICRA)*, pp 3972–3976
18. King CH, Culjat MO, Franco ML, Lewis CE, Dutton EP, Grundfest WS, Bisley JW (2009) Tactile feedback induces reduced grasping force in robot-assisted surgery. *IEEE Trans Haptics* 2(2):103–110
19. King CH, Chen TL, Jain A, Kemp CC (2010) Towards an assistive robot that autonomously performs bed baths for patient hygiene. In: *Proceedings of the 2010 IEEE/RSJ international conference on intelligent robots and systems (IROS)*, pp 319–324
20. Koga H, Usuda Y, Matsuno M, Ogura Y, Ishii H, Solis J, Takanishi A, Katsumata A (2007) Development of oral rehabilitation robot for massage therapy. In: *Proceedings of the 6th international special topic conference on information technology applications in biomedicine (ITAB)*, pp 111–114
21. Kulic D, Croft E (2007) Pre-collision safety strategies for human-robot interaction. *Auton Robots* 22(2):149–164
22. Lacey C, Finkelstein M, Thygeson M (2008) The impact of positioning on fear during immunizations: supine versus sitting up. *J Pediatr Nurs* 23(3):195–200
23. Lang P (1995) The emotion probe: studies of motivation and attention. *Am Psychol* 50(5):372
24. Lang P, Bradley MM, Cuthbert BN (1997) *International affective picture system (IAPS): technical manual and affective ratings*. NIMH Center for the Study of Emotion and Attention
25. Mandryk R, Atkins M (2007) A fuzzy physiological approach for continuously modeling emotion during interaction with play technologies. *Int J Hum-Comput Stud* 65(4):329–347
26. Maselli M, Altrocchi J (1969) Attribution of intent. *Psychol Bull* 71(6):445
27. McCann K, McKenna H (1993) An examination of touch between nurses and elderly patients in a continuing care setting in Northern Ireland. *J Adv Nurs* 18(5):838–846
28. Meisner E, Isler V, Trinkle J (2008) Controller design for human-robot interaction. *Auton Robots* 24:123–134
29. Morrison I, Löken LS, Olausson H (2010) The skin as a social organ. *Exp Brain Res* 204(3):305–314
30. Mukai T, Hirano S, Nakashima H, Kato Y, Sakaida Y, Guo S, Hosoe S (2010) Development of a nursing-care assistant robot RIBA that can lift a human in its arms. In: *Proceedings of the 2010 IEEE/RSJ international conference on intelligent robots and systems (IROS)*, Taipei, Taiwan, pp 5996–6001
31. Nakagawa K, Shiomi M, Shinozawa K, Matsumura R, Ishiguro H, Hagita N (2011) Effect of robot's active touch on people's motivation. In: *Proceedings of the 6th international conference on human-robot interaction*, pp 465–472
32. Nomura T, Kanda T, Suzuki T (2006) Experimental investigation into influence of negative attitudes toward robots on human-robot interaction. *AI Soc* 20(2):138–150
33. Premack D, Premack A (1995) Origins of human social competence. *Cogn Neurosci* 10:205–218
34. Riskind J, Gotay C (1982) Physical posture: could it have regulatory or feedback effects on motivation and emotion? *Motiv Emot* 6(3):273–298
35. Rosenthal-von der Pütten AM, Krämer NC, Hoffmann L, Sobieraj S, Eimler SC (2013) An experimental study on emotional reactions towards a robot. *Int J Soc Robot* 5(1):17–34
36. Saerbeck M, Bartneck C (2010) Perception of affect elicited by robot motion. In: *Proceedings of the 5th ACM/IEEE international conference on human-robot interaction (HRI)*. IEEE, New York, pp 53–60
37. Shen Q, Kose-Bagci H, Saunders J, Dautenhahn K (2011) The impact of participants' beliefs on motor interference and motor coordination in human-humanoid interaction. *IEEE Trans Auton Ment Dev* 99:1
38. Shiomi M, Kanda T, Ishiguro H, Hagita N (2006) Interactive humanoid robots for a science museum. In: *Proceedings of the 1st ACM SIGCHI/SIGART conference on human-robot interaction*, pp 305–312
39. Swangnetr M, Zhu B, Kaber D, Taylor K (2010) Meta-analysis of user age and service robot configuration effects on human-robot interaction in a healthcare application. In: *2010 AAAI fall symposium series*
40. Tsui KM, Desai M, Yanco HA, Cramer H, Kemper N (2010) Using the “negative attitude toward robots scale” with telepresence robots. In: *Proceedings of the 10th performance metrics for intelligent systems workshop*, pp 243–250
41. Tsumaki Y, Kon T, Suginuma A, Imada K, Sekiguchi A, Nenchev DN, Nakano H, Hanada K (2008) Development of a skincare robot. In: *2008 proceedings of the IEEE international conference on robotics and automation (ICRA)*, pp 2963–2968
42. Wagner A, Arkin R (2011) Acting deceptively: providing robots with the capacity for deception. *Int J Soc Robot* 3(1):5–26
43. Watson D, Clark LA, Tellegen A (1988) Development and validation of brief measures of positive and negative affect: the PANAS scales. *J Pers Soc Psychol* 54(6):1063
44. Yohanan S, MacLean KE (2012) The role of affective touch in human-robot interaction: human intent and expectations in touching the haptic creature. *Int J Soc Robot* 4(2):163–180
45. Zoghbi S, Croft E, Kulic D, Van der Loos M (2009) Evaluation of affective state estimations using an on-line reporting device during human-robot interactions. In: *Proceedings of the 2009 IEEE/RSJ international conference on intelligent robots and systems (IROS)*, pp 3742–3749

Tiffany L. Chen is a Robotics PhD Student at the Georgia Institute of Technology. She earned B.S. and M.S.E. degrees in Biomedical Engineering from Johns Hopkins University in 2007 and 2008, respectively. Her research has been published in the areas of human-robot interaction, assistive and rehabilitation robotics, and surgical training.

Chih-Hung Aaron King contributed to this work as a postdoctoral researcher in the Healthcare Robotics Lab at the Georgia Institute of Technology. He is currently the R&D Manager of DTI Robotics US Inc., the US subsidiary of Danish Technological Institute. He received his B.S. in Electrical Engineering in 2000, M.S. in Biomedical Engineering in 2006, and Ph.D. in Biomedical Engineering in 2008 from University of California, Los Angeles. His research has been in haptic sensing and actuation, dexterous mobile manipulation, human-robot interaction, and flexible industrial robots.

Andrea L. Thomaz is an Assistant Professor of Interactive Computing at the Georgia Institute of Technology. She directs the Socially Intelligent Machines lab, which is affiliated with the Robotics and Intelligent Machines (RIM) Center and with the Graphics Visualization and Usability (GVU) Center. She earned a B.S. in Electrical and Computer Engineering from the University of Texas at Austin in 1999, and Sc.M. and Ph.D. degrees from MIT in 2002 and 2006. Dr. Thomaz has published in the areas of Artificial Intelligence, Robotics, Human-Robot Interaction, and Human-Computer Interaction.

Charles C. Kemp is an Associate Professor at the Georgia Institute of Technology in the Department of Biomedical Engineering, and has adjunct appointments in the School of Interactive Computing and the School of Electrical and Computer Engineering. He earned a doctorate in Electrical Engineering and Computer Science (2005), an MEng, and BS from MIT. In 2007, he founded the Healthcare Robotics Lab at Georgia Tech (<http://healthcare-robotics.com>), which focuses on mobile manipulation and human-robot interaction.