# Attribution of Blame in a Human-Robot Interaction Scenario

Federico Scholcover
*University of Central Florida*

# ATTRIBUTIONS OF BLAME IN A HUMAN-ROBOT INTERACTION SCENARIO

by

FEDERICO SCHOLCOVER

A thesis submitted in partial fulfillment of the requirements
for the Honors in the Major Program in Psychology
in the College of Sciences
and in The Burnett Honors College
at the University of Central Florida
Orlando, Florida

Spring Term 2014

Thesis Chair: Dr. Valerie Sims

# Abstract

This thesis worked towards answering the following question: Where, if at all, do the beliefs and behaviors associated with interacting with a nonhuman agent deviate from how we treat a human? This was done by exploring the inter-related fields of Human-Computer and Human-Robot Interaction in the literature review, viewing them through the theoretical lens of anthropomorphism. A study was performed which looked at how 104 participants would attribute blame in a robotic surgery scenario, as detailed in a vignette. A majority of results were statistically non-significant, however, some results emerged which may imply a diffusion of responsibility in human-robot collaboration scenarios.

# Dedications

For Gilbert.

Who's a good boy?

## Acknowledgements

To begin with, I would like to acknowledge and thank my committee: Dr. Sims, Dr. Chin, Dr. Maboudou, and Dr. Sinatra. Thank you for dealing with me these last few months and my prolific talent for procrastination. I would also like to thank every single member of the Applied Cognition And Technology lab. Without the support and friendship you have all shown me, I doubt I would be where I am. I would start listing out names and thanking you each individually, but I don't think the acknowledgment section is supposed to be longer than my thesis. I know I've already mentioned their names, but I would like to especially thank Dr. Sims and Dr. Chin. Our weekly lunch meetings are always one of the highlights of my week, and the two of you really exemplify what it means to be a mentor. Dr. Sims, I know I always ask if I've thanked you yet, for one thing or another. I think it may be because I always feel like I haven't been able to express it enough! So, for one more time, thank you.

# Table of Contents

## Table of Figures

# Introduction

This thesis will explore human-robot interaction through the study of anthropomorphism, using what we know of ourselves to understand other entities, as well as the link with the field of human-computer interaction. There is a running theme within the literature: In an overwhelming amount of studies, it does not matter if we are speaking about humans, robots, computers, pets, rocks, or pet rocks; Given that they show sufficient social cues, we utilize the same underlying system to understand of all of those entities. The literature review and experiment will look at answering the following question: Where, if at all, do the beliefs and behaviors associated with interacting with a nonhuman agent deviate from how we treat a human? This question will be applied in the context of robotic surgery through an experiment which will ask the participants to attribute blame for post-operative complications which may or may not have occurred during surgery. Future directions will also be discussed.

# Literature Review

Before beginning, it may be of use to establish a working definition of some commonly used terms, as their definition may vary within the literature. For example, Nowak and Biocca (2003) define an agent as a computerized entity whose actions are controlled by the computer while Epley, Waytz, and Cacioppo (2007) define an agent as any being, including animals and devices, which acts with "apparent independence." As used in this thesis, an agent is defined as any being capable of at least some level of self-determination through the ability to react to its environment, whether physical or virtual, in a controlled manner. Higher levels of agency imply a greater ability to exert control over the environment. A computer is an agent and a human is an agent, however, a human has a higher degree of agency as it is able to exert more control over its environment. An avatar is the embodiment of such an agent within a virtual environment.

## Anthropomorphism

Anthropomorphism is defined as the tendency to attribute human-like characteristics to nonhuman agents. These human-like characteristics include such things as intentionality, motivations, thoughts, and emotions (Epley, Waytz, & Cacioppo, 2007). Another source defined anthropomorphism as the rationalization of behavior by superimposing aspects of the human observer (Duffy, 2003). Both definitions work towards the same idea: anthropomorphism serves as a medium from which to understand the behavior of a non-human entity (and arguably a human, as well). It performs this action by creating inferences of the anthropomorphized entity's behavior using an observer's knowledge of their self. A robotic vacuum, such as a Roomba, is a great example of a possibly anthropomorphized agent. A robotic vacuum is capable of acting with a degree of independence. It rolls around a floor, vacuuming as it detects dirt with minimal input from a user. In a popular television show, a character has even attributed a mental state to the robotic vacuum, stating “Oh, it’s hungry,” as the robot beeped, indicating that it would begin to vacuum (Hurwitz, 2005).

Epley, Waytz, and Cacioppo (2007) proposed a model to explain what causes a human to anthropomorphize an agent. The model contains three psychological determinants: Elicited agent knowledge, effectance motivation, and sociality motivation. They define elicited agent knowledge as the act of using the knowledge one has about themselves to make inferences about the state of an agent, effectance motivation as the need or motivation to understand the agent, and they described sociality motivation as the innate need each human has to make social connections.

Eyssel, Kuchenbrandt, and Bobinger (2011) explored this concept by having participants watch a video and read a vignette about a robot. Participants were then given scales which assessed how “human” these robots appeared. The researchers made the assertion that the more humanness a participant attributed to the robot, the more they were anthropomorphizing the robot. The study looked at two factors from the Epley, Waytz and Cacioppo model: One factor involved influencing sociality motivation. Half of the participants were informed that they would later be interacting with the robot, the other half were not. The other factor worked to influence effectance motivation: half of the participants read a vignette designed to make them believe the robot was unpredictable, while the other half were lead to believe it was predictable. Participants who believed they would be interacting with the robot, looking at sociality motivation, rated the robot as more human. These effects, however, were not shown for predictability. However, predictability did interact with sociality motivation, giving support for effectance motivation. Participants who believed the robot was unpredictable and were informed that they would be interacting with the robot gave the highest rating of humanness to the robot. It was as if they were working harder to understand the unpredictable robot, knowing that they would have to interact with it, causing higher levels of anthropomorphism.

**Generalizability of Social Behaviors**

A key concept in the study of Human-Computer interaction is known as the Computers as Social Actors hypothesis, which expands from the realm anthropomorphism. It states that if you were to take any social science finding which describes how humans interact with other humans, you could also correctly apply those same rules when humans interact with a computer, given that the computer exhibits sufficient social cues or rather, almost any characteristic which may be thought of as “human” such as the ability to produce speech (Nass, Steuer, & Tauber, 1994). This suggests that not only do humans use similar social behaviors when interacting with other humans or non-human entities, but that there is a very low threshold of social cues that needs to be surpassed before we begin to use these similar social behaviors.

In a study looking at a user’s sense of co-presence, the feeling of being in the company of others, found that it did not matter whether participants thought an avatar they were interacting with was controlled by a human or a computer, the participants had no difference in their sense of copresence (Nowak & Biocca, 2003). Another study found that the way adults and children respond to avatars, based upon an avatar’s attractiveness parallels how they respond to another human based upon the human’s attractiveness. The more attractive they rate the avatar, the more positive they rate the avatar along a variety of dimensions (Principe, Langlois, 2013). These two studies help to illustrate that, not only do the social science findings which describe human-

human interaction apply to human-nonhuman interaction, but that they will apply these systems mindlessly.

Studies have also shown that these results expand to human-robot interaction. A paper by Eyssel and Kuchenbrandt (2011), for example, found that the concept of grouping bias, preferential behavior towards those of one's own group, extended towards robots. They found that anthropomorphism is a function of robot group membership. That is, participants generally rated the in-group robot much more favorably than the out-group robot on measures of warmth, design preference, psychological closeness, willingness to live and talk to the robot, as well as attributing more mental states to the robot, which parallels what you find in classic grouping bias papers.

**Inconsistencies in the Anthropomorphism Literature**

Embodied Conversational Agents (ECAs) can be thought of as a computer controlled, human-like, avatars embedded in a computer interface capable of interacting with a user. One of the design goals in the inclusion of these agents is to improve task performance, often in reading passages. They attempt to do this by increasing task engagement, while avoiding the danger of becoming too distracting (Yee, Bailenson, & Rickertsen, 2007).

A meta-analysis looking at the effects that ECAs have on user interfaces, found that having these agents produced more positive social interactions than not having them, and increases in the realism of these agents was positively correlated with positive social interactions as well. However, the correlation in positive social interaction with realism was only seen when looking at subjective measures; when looking at behavioral measures, there was no correlation with increased realism. That is to say: Participants would alter their pattern of reporting based on differences in realism, but they would not actually behave any differently. However, they warn that it may be due to a lack of measurement sensitivity rather than something else. Interestingly, the meta-analysis goes on to show that the effect size for many of the comparisons looking at the inclusion/exclusion of an agent were much higher than those comparing low and high realism (Yee, et al., 2007). Gong (2008) posited that degrees of anthropomorphism are not operationally well defined in the body of anthropomorphism literature. What is considered highly anthropomorphic in one study may not be in another, which may partially explain the issue of small effect sizes found for the low/high realism comparison found by Yee.

However, the results found in Yee may imply that, on a behavioral level, anthropomorphism does not happen in degrees. A human is either anthropomorphizing something or not. This creates an interesting point, when looking at the Eyssel and Kuchenbrandt (2011) study discussed on the previous page, as they found that anthropomorphism increased when a robot was part of an in-group rather than out-group. However, it is possible that they were studying a related concept as they defined anthropomorphism as a function of such things as psychological closeness. Participants may have just been applying the appropriate social behavior for the situation, not anthropomorphizing the in-group robot any more or less than the other robot. This brings back the argument laid out by Gong (2008): “degrees of anthropomorphism” is a poorly defined concept.

In a study looking at participants’ reaction to a robot’s proxemic behaviors, how the robot comes into proximity of a human (Syrdal, Koay, Walters, & Dautenhahn, 2007), researchers found that women generally allowed a robot to approach them much closer from the front than male participants, but that both males and females preferred robots to approach from the sides, as they reported that having the robot approach from the sides was less confrontational. They also found that participants high in extraversion would allow the robot to approach closer when a human was in control of the robot than when the robot was in control, but that there were no differences in the control conditions for participants low in extraversion. The results were also paralleled when measuring conscientiousness, as the participants scoring low in conscientiousness would allow the robot to approach closer when a human was in control of the robot than when the robot was in control, with no differences between the human/robot control conditions for those high in conscientiousness (Syrdal, Koay, Walters, & Dautenhahn, 2007). These differences which emerge when a robot is in control versus a human being in control of a robot may imply that there is some underlying difference in social behaviors depending on what is in control.

Kim and Hinds (2006) studied attributions of blame and credit in a human-robot interaction scenario. Participants were tasked with assembling an object, which was then transported by a robot to another participant, who further assembled the object. The participants were lead to believe the robot was automated. This was done using a “Wizard of Oz” approach. The researchers manipulated how automatic the robot appeared to be as well as how transparent its actions were.  To create the transparency manipulation, the robot would spontaneously spin three times. In the low transparency condition, the robot would do this without an explanation. However, in the high transparency condition, the robot would announce “I have recalibrated my sensors.”  They altered how automatic it was perceived by participants by either having the robot move once an object was readied for transportation or by waiting for the participant to tell the robot to move. They found that how automatic the robot became had a positive correlation with how much blame the robot received for an unsuccessful task completion and a negative correlation with how much participants would blame themselves. The more automatic it became, the more likely participants were to blame the robot and the less likely they were to blame themselves. Interestingly, this effect did not extend towards crediting the robot. There was no effect for higher levels of automaticity on how much participants credited the robot for a successful task completion. The researchers did not find any significant effects for transparency as it relates to the specific participant and the robot. However, the participants did both blame and credit their coworkers less when in the high transparency condition. The fact that participants would blame a robot more the more automatic it became may imply an increase in anthropomorphism. Relating this study back to the model on anthropomorphism by Epley, Waytz, and Caccioppo (2007), an increase in how automatic the robot became may be effecting effectance motivation, which was earlier defined as the need or motivation to understand the agent. While it is arguable that the transparency manipulation was more akin to increasing effectance motivation, the more automatic a robot becomes, it may lead the participant to start asking which cues the robot uses to begin moving, increasing effectance motivation.

## Robotic Surgery

Robotic assisted surgery involves the use of a robotic intermediary between the patient and the surgeon. Robots were initially introduced into the operating room to function as camera holders, receiving FDA approval for their use in 1994 (Ballantyne, 2002). However, they have moved to the stage where they can perform minimally invasive surgery. This field is seeing rapid expansion, and, despite being in its relative infancy, it has already proved its worth in the realm of minimally invasive surgeries. Certain surgical systems like the da Vinci and Zeus systems filter out a surgeon's hand tremors and, due to their multiple degrees of freedom (mechanical degrees of freedom, not statistical), they are capable of reaching areas a human hand might not reach as easily (Lanfranco, Castellanos, Desai, & Meyers, 2004). However, the nature of these robots means that the surgeon may be sitting six or more feet away from the patient, tele-operating the robot, viewing the surgery through monitors. This may lead to the surgeon feeling like they have less control over the surgery (Gerhardus, 2003). The teleoperation of these robots implies a possible link with research on avatars, as, in a sense, the robot is the physical embodiment of the surgeon's avatar. Focusing on the previously discussed study by Nowak and Biocca (2003), participants perceived no differences in presence and copresence between a human controlled and a computer controlled avatar. This may imply that participants will not perceive any differences between an automated surgical robot and a human controlled surgical robot.

**Research Questions and Hypothesis**

The theme throughout this literature review has been that, on many dimensions, we treat humans no different than any other types of agents. However, there are some areas which differences do occur. Looking particularly at the Kim and Hinds (2006) study, participants would blame the robot more as it became more automatic. However, the researches explored this within the realm of a collaborative task: The participant was a member of the ongoing event. What would happen if we explored these attributions of blame through the perspective of a third party? What if the human and robot were performing the same task collaboratively, rather than a different but collaborative task?

Looking at the Syrdal, Koay, Walters, & Dautenhahn (2007) study, participants would react differently to a robot, depending on if the robot was in control of itself or if a human was in control of the robot. What would happen if a robot was somehow controlling a human?

The questions which emerge from the two studies were used to construct the manipulations for the experiment. One manipulation, which will later be referred to as the Incision manipulation, will look to see if the differences in blame found in Kim and Hinds were because a robot was performing a task instead of a human. The other, which will later be referred to as the Decision manipulation, will look to see if the differences of who is in control found in Syrdal, et. al. generalize to artificial entities, rather than humans, being in control and, with the use of the other manipulation, humans being controlled by either another human or an artificial entity. See table below for further clarification.

| | |
|---|---|
| Human in control<br><br>Human performing task | Artificial Entity in control<br><br>Artificial Entity performing task |
| Artificial Entity in control<br><br>Human performing task | Human in control<br><br>Artificial Entity performing task |

**Figure 1 Table which describes possible combinations of control and task performance.**

**Results from Syrdal et. al. suggest that conditions in the same row will be similar to each other, but different from conditions in different rows.**

This leads to the research hypothesis: There is some underlying effect, in which differences in the type of entity which is in control of a task will lead to different attributions of blame from the perspective of a human observer. This may also generalize to an entity which is performing, rather than controlling, the actual task.

# Method

*Participants*

A total of 268 survey responses were collected using an online survey system. However, 164 responses were omitted from the analysis. The final sample for analysis included 104 undergraduate male (N=36) and female (N=68) students.

Of the 164 responses that were omitted, 50 survey responses came from participants who took the survey more than once and who were exposed to an experimental condition each time. An additional 114 responses were removed due to failure to follow instructions, including failure to complete the survey and incorrectly answering questions which assessed one of two things: Whether the participant had paid attention to the vignette or if the participant was even paying attention to the question being asked. Some participants incorrectly responded to the following "question": "For this question, please select answer choice 1."

*Procedure*

The participants were asked to read an excerpt of a fake medical transcript, adapted from multiple medical transcripts found online which utilized the Da Vinci Surgical System (MTSamples.com, n.d.). The experimental transcript the participants read outlined a medical diagnosis that required surgery, the accompanying surgical procedure, as well as the patient's post-operative outcome. The patient was said to have "experienced post-operative complications" for all conditions. The participants were then asked to fill out a variety of questions which assessed when, where, how much, and to whom they attributed blame for the post-operative complications.

The experiment itself is a 3x2 factorial design, where certain entities within the medical transcript were changed between participants. The entity(s) performing the actual surgical procedure, referred to as the "Incision" manipulation, was either a 5th year surgical resident, a fully automated Asimov Surgical System, or a 5th year surgical resident controlling an Asimov Surgical System[1]. The use of a surgical resident was chosen, rather than an actual attending surgeon, as they have received sufficient medical training to perform surgery, but they are still inexperienced enough that it is not unusual to have a more experienced surgeon in the room supervising. The entity performing the diagnosis as well as supervising the surgery, referred to as the "Decision" manipulation, was either an attending surgeon or the fictional Asimov Surgical A.I. [2] The Decision manipulation appears twice within the vignette to emphasize the fact that the Decision entity is not actually performing any tasks, simply supervising. See figure below for a summary of the manipulations.

| **Incision** / **Decision** | **Resident** | **Resident/Robot** | **Robot** |
|---|---|---|---|
| **Surgeon** | Surgeon supervises. Resident performs surgery | Surgeon supervises. Resident operates robot to perform surgery. | Surgeon supervises. Automated robot performs surgery. |
| **A.I.** | A.I. supervises. Resident performs surgery | A.I. supervises. Resident operates robot to perform surgery. | A.I. supervises. Automated robot performs surgery. |

**Figure 2 Summary of events occurring within experimental manipulations.**

*Materials*

[1] From now on, referred to as the resident, the robot, and the resident/robot condition respectively

[2] From now on, referred to as the surgeon and the A.I. condition, respectively.

A sample vignette is included on the next page, with explanatory footnotes. A copy of each vignette is available in the appendix. The questions participants were asked to answer are listed on the page proceeding the sample vignette.

Please read the following brief excerpt of a transcript of a medical procedure carefully.
You ***WILL*** be asked questions about what you read:

**Key Phrases:**

*Hematuria*: Blood in urine
*Dysuria*: Painful Urination
*PSA*: Prostate-Specific Antigen
*Adenocarcinoma*: A type of cancer
*Retropubic Prostatectomy*: A surgical procedure requiring anesthesia

[1]Asimov Surgical System: A robotic system with five surgical arms, each capable of performing different surgical tasks, including incisions and endoscopies.
[2]Asimov Surgical A.I.: An artificial intelligence developed by the Asimov Corporation, capable of performing analytical tasks, such as diagnostics and surgical supervision.
Asimov Corporation: A corporation specializing in the development and manufacturing of high-tech surgical tools.

**Transcript:**

**Brief History of the Patient:** The patient is a 45-year-old male with a history of urological issues, including a previous hematuria caused by right renal stones. He reported, to the intake nurse, a return of the hematuria, but with incidences of dysuria and acute pain in his pelvic bone. Blood tests returned with an abnormally high PSA level. The patient was given a preoperative diagnosis of adenocarcinoma of the prostate by the DECISION[3]. Upon consideration of various treatment modalities and their associated preoperative prognoses, the patient decided to undergo a retropubic prostatectomy, in order to remove the adenocarcinoma.

**Procedure:** The patient was taken into the operating room, and successfully given a general anesthetic. As per the patient's request, the surgery was performed by a INCISION[4] under the supervision of the DECISION[5]. The surgery was performed to standard.

**Postoperative Status:** Post-surgical blood tests have indicated a return to normal PSA levels, as well as no signs of the adenocarcinoma. However, due to the nature of the procedure, the patient has experienced postoperative complications.

[1] Included only when the incision factor level included the Asimov Surgical System
[2] Included for the Asimov Surgical A.I. decision factor level.
[3] Decision Condition: Will either say "attending surgeon" or "Asimov Surgical A.I."
[4] Incision Condition: Will say either "5th year surgical resident," "5th year surgical resident controlling an Asimov Surgical System," or "fully automated Asimov Surgical System."
[5] Will always match Decision Condition.

| Vignette Questions | |
|---|---|
| Perceived Chance of Error[1] | |
| | How likely is it that the post-operative complications were due to an error? |
| | How likely is it that the post-operative complications were due to random chance? |
| Temporal Placement of Error[1] | |
| | How likely is it that there was an error in the diagnosis prior to the surgery was performed? |
| | How likely is that the correct surgical procedure was performed? |
| | How likely is that an error occurred after the surgery? |
| Type of Error[1] | |
| | How likely is it that a mental error occurred during surgery? |
| | How likely is it that a physical error occurred during surgery? |
| Entity at Fault [1] | |
| | How likely are you to find the hospital at fault for the post-operative complications? |
| | How likely are you to find Asimov Corporation at fault for the post-operative complications? |
| | How likely are you to find the diagnostic team, as a whole, at fault for the post-operative complications? |
| | How likely are you to find the surgical team, as a whole, at fault for the post-operative complications? |
| | How likely are you to find the patient at fault for the post-operative complications? |
| | How likely are you to find the intake nurse at fault for the post-operative complications? |
| Compensation | |
| | How much compensation does the patient deserve for the post-operative complications?[2] |
| | Assume the patient will be receiving compensation for the resulting medical expenses. How much should the patient be compensated for the resulting medical expenses and general damages? Please enter a number between $0 and $100,000.[3] |

[1] 5-point Likert Scale: (Extremely Unlikely, Less Likely Than Not, Somewhat Likely, More Likely Than Not, Extremely Likely)

[2] 5-point Likert Scale: (The patient deserves no compensation, The patient deserves a small compensation, The patient deserves a moderately sized compensation, The patient deserves a large compensation, The patient deserves as large of a compensation as he can get.)

[3] Only values between 0 and 100,000 accepted.

# Results

Unless otherwise specified, all analyses were run as a two way, full factorial, between-subjects ANOVA with the factors: Incision and Decision. All tests are performed at the $\alpha$=.05 significance level.

## Perceived Chance of Error

*How likely is it that the post-operative complications were due to an error?*

Results indicate that there was a nonsignificant interaction, $F(2,98)=.145$, $p=.865$, and a nonsignificant main effect for Decision, $F(1,98)=.139$, $p=.710$. However, there was a marginally significant main effect for Incision $F(2,98)= 2.848$, $p=.063$. A Tukey Honest Significant Differences was run to identify which means within the Incision manipulation were approaching significantly different. The Resident/Robot condition ($M$=2.086, $SD$=.781) was marginally significantly different from the Robot condition ($M$=2.487, $SD$=.823), $p=.081$. Means for the Resident condition ($M$=2.467, $SD$=.730) were not significantly different from either group.

*How likely is it that the post-operative complications were due to random chance?*

Results indicate no significant interaction, $F(2,98)=.323$, $p=.725$, and no significant main effects for Incision, $F(2,98)=.231$, $p=.808$, or for Decision, $F(1,98)=.309$, $p=.579$. Participants, in general, reported that it was “somewhat likely” that the post-operative complications were due to random chance ($M$=3.135, $SD$=1.0054).

**Temporal Placement of Error**

*How likely is it that there was an error in the diagnosis prior to the surgery was performed?*

Results indicate no significant interaction, $F(2,98)=.323$, $p=.725$, and no main effects for Incision, $F(2,98)=.200$, $p=.819$, or for Decision, $F(1,98)=1.473$, $p=.228$. Participants, in general, reported that it was "less likely than not" that there was an error prior to the diagnosis being performed ($M=2.183$, $SD=.707$).

*How likely is it that the correct surgical procedure was performed?*

Results indicate that there was a marginally significant interaction $F(2,98)=1.408$, $p=.074$. There was no main effect for Incision, $F(2,98)=1.577$, $p=.212$, or for Decision $F(1,98)$-1.791, $p=.184$. Levene's Test was nonsignificant as well, $F(5,98)=.894$, $p=.488$. See the table on the next page for means and standard deviations by condition.

| Incision/Decision | A.I. | Surgeon | Total |
|---|---|---|---|
| Resident | *M*=4.36, *SD*=.63 | *M*=3.69, *SD*=.704 | *M*=4.00, *SD*=.747 |
| Resident/Robot | *M*=4.07, *SD*=.917 | *M*=4.24, *SD*=.436 | *M*=4.17, *SD*=.664 |
| Robot | *M*=3.94, *SD*=.873 | *M*=3.81, *SD*=.750 | *M*=3.87, *SD*=.801 |
| Total | *M*=4.11, *SD*=.823 | *M*=3.93, *SD*=.672 | *M*=4.01, *SD*=.744 |

**Figure 3 Means and standard deviations for the question "How likely is it that the correct surgical procedure was performed?"**

*How likely is that an error occurred after the surgery?*

Results indicate no significant interaction, $F(2,98)=1.09$, $p=.340$, and no significant main effects for Incision, $F(2,98)=1.243$, $p=.293$, or Decision, $F(1,98)=.655$, $p=.420$. Participants, in general, reported that it was "somewhat likely" that an error occurred after the surgery ($M=2.80$, $SD=.781$).

## Type of Error

*How likely is it that a mental error occurred during surgery?*

Results indicate and no significant interaction, $F(2,98)=.549$, $p=.580$ and no significant main effects for Incision, $F(2,98)=1.230$, $p=.297$, or Decision, $F(1,98)=.367$, $p=.546$. Participants, in general, reported that it was “less likely than not” that a mental error occurred during surgery ($M=2.08$, $SD=.784$).

*How likely is it that a physical error occurred during surgery?*

Results indicate no significant interaction, $F(2,98)=.245$, $p=.783$, and no significant main effects for Incision, $F(2,98)=1.015$, $p=.366$, or Decision, $F(1,98)=.126$, $p=.723$. Participants, in general, reported that it was “less likely than not” that a physical error occurred during surgery ($M=2.47$, $SD=.955$).

## Entity At Fault

*How likely are you to find the hospital at fault for the post-operative complications?*

Results indicate no significant interaction, $F(2,98)=1.124$, $p=.329$, and no significant main effects for Incision, $F(2,98)=1.15$, $p=.321$, or Decision, $F(1,98)=.017$, $p=.898$. Participants, in general, reported that it was “less likely than not” that the hospital was at fault for the post-operative complications ($M=2.375$, $SD=.957$).

*How likely are you to find Asimov Corporation at fault for the post-operative complications?*

Results indicate no significant interaction, $F(2,98)=.701$, $p=.499$ and no significant main effects for Incision, $F(2,98)=1.369$, $p=.259$, or Decision, $F(1,98)=.737$, $p=.393$. Participants, in general, reported that it was “less likely than not” that the Asimov Corporation was at fault for the post-operative complications ($M=2.44$, $SD=.964$).

*How likely are you to find the diagnostic team, as a whole, at fault for the post-operative complications?*

Results indicate no significant interaction, $F(2,98)=2.079$, $p=.131$, and no significant main effects for Incision, $F(2,98)=.617$, $p=.542$, or Decision, $F(1,98)=.711$, $p=.401$. Participants, in general, reported that it was “less likely than not” that the Diagnostic Team was at fault for the post-operative complications ($M=2.279$, $SD=.939$).

*How likely are you to find the surgical team, as a whole, at fault for the post-operative complications?*

Results indicate no significant interaction, $F(2,98)=1.034$, $p=.360$, and no significant main effects for Incision, $F(2,98)=.034$, $p=.967$, or Decision, $F(1,98)=.334$, $p=.565$. Participants, in general, reported that it was “somewhat likely” that the diagnostic team was at fault for the post-operative complications ($M=2.67$, $SD=.990$)

*How likely are you to find the patient at fault for the post-operative complications?*

Results indicate no significant interaction, $F(2,98)=.153$, $p=.859$, and no significant main effects for Incision, $F(2,98)=.102$, $p=.903$, or Decision, $F(1,98)=.145$, $p=.704$. Participants, in general, reported that it was “less likely than not” that the patient was at fault for the post-operative complications (*M*=2.067, *SD*=1.007).

*How likely are you to find the intake nurse at fault for the post-operative complications?*

Results indicate no significant interaction, $F(2,98)=.937$, $p=.395$, and no significant main effect for Incision, $F(2,98)=.551$, $p=.578$. However, there is a significant main effect for Decision, $F(1,98)=4.089$, $p=.046$. While both factor levels of the Decision condition reported that it was “less likely than not” that the intake nurse was at fault for the post-operative complications (*M*=1.88, *SD*=.780), participants in the A.I. condition reported a lower mean (*M*=1.713, *SD*=.834) than did the participants in the Surgeon condition (*M*=2.021, *SD*=.713). Levene’s Test came out nonsignificant as well, $F(5,98)=.668$, $p=.649$.

*Summary of Results for Entity Questions*

The mean scores for each of the six entities are listed in the table below, along with the upper and lower bounds for the 95% confidence intervals generated from the means. All upper bounds were below the value of 3, indicating that the means for all of the entities were significantly less than the midpoint value for blame; that the entity was at fault for the post-operative conditions. This indicates that participants generally did not find any of the entities to be at fault for the post-operative complications.

| Entity | Mean | Lower Bound | Upper Bound |
|---|---|---|---|
| Hospital | 2.38 | 2.19 | 2.56 |
| Asimov Corporation | 2.44 | 2.25 | 2.63 |

| | | | |
|---|---|---|---|
| Diagnostic Team | 2.28 | 2.10 | 2.46 |
| Surgical Team | 2.67 | 2.48 | 2.87 |
| Patient | 2.07 | 1.87 | 2.26 |
| Intake Nurse | 1.88 | 1.73 | 2.04 |

**Figure 4 Summary of Results for Entity Questions**

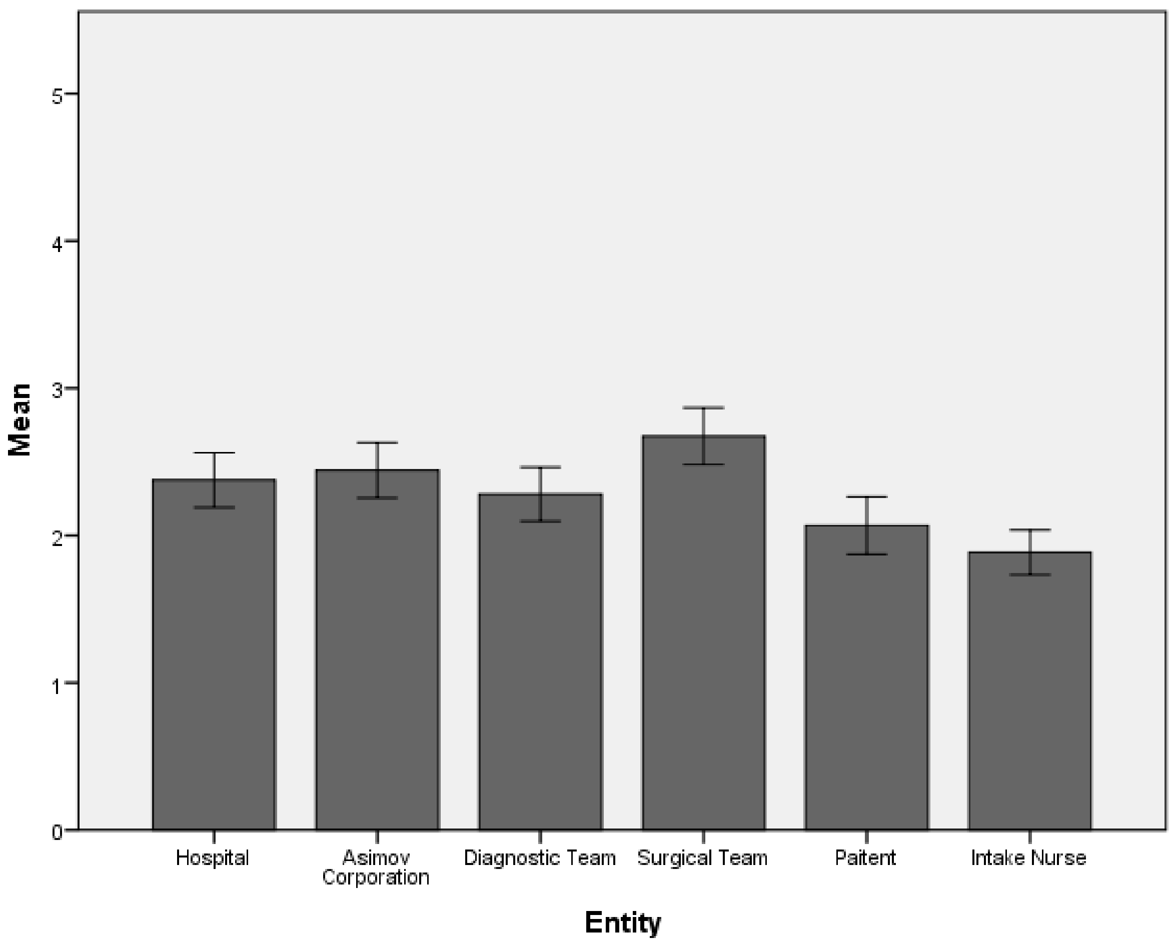


**Figure 5 Graphic of Results for Entity Questions**

**Compensation**

*How much compensation does the patient deserve for the post-operative complications?*

Results indicate no significant interaction $F(2,98)=.689$, $p=.505$, and no significant main effect for Incision, $F(2,98)=.658$, $p=.520$. However, there is a significant main effect for Decision, $F(1,98)=4.636$, $p=.034$. Participants in the A.I. condition reported that "the patient deserves a moderately sized compensation." ($M=2.732$, $SD=1.06$) while participants in the Surgeon condition reported that "the patient deserves a small compensation." ($M=2.299$, $SD=.973$). Levene's Test came out nonsignificant as well, $F(5,98)=.878$, $p=.499$.

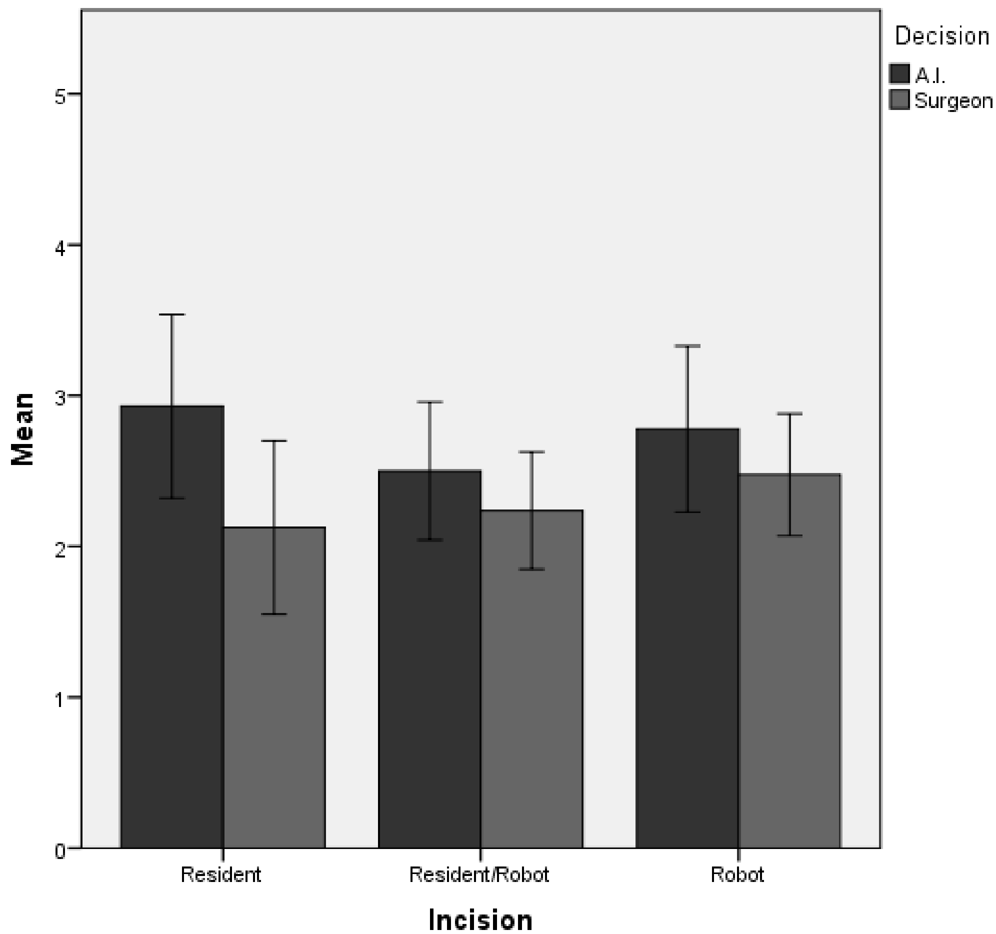


**Figure 6 Graphic of Results for "How much compensation does the patient deserve for the post-operative complications?"**

*Assume the patient will be receiving compensation for the resulting medical expenses. How much should the patient be compensated for the resulting medical expenses and general damages? Please enter a number between $0 and $100,000.*

Results indicate was no significant interaction $F(2,98)=.793$, $p=.455$ and no significant main effect for Decision, $F(1,98)=.324$, $p=.571$. There is a significant main effect for Incision, $F(2,98)=5.800$, $p=.004$. However, Levene's Test came out significant, $F(5,98)=2.325$, $p=.048$, indicating that the data is heteroscedastic, violating an underlying assumption of the ANOVA model

A Kruskal-Wallis One Way ANOVA was performed to look for differences in the medians of the Incision conditions. There was a significant difference in the medians of the Incision conditions, $\chi^2(2)$=.11.580, $p$=.003. The Resident/Robot condition (*Md*=10000) had a lower median than both the Resident (*Md*=25000) and Robot (*Md*=25000), which each had the same medians. Looking at means, the pattern was still maintained: Participants in the Resident/Robot condition (*M*=14020, *SD*=18302.73) reported that the patient should be compensated less than was reported by the participants in both the Resident (*M*=33267, *SD*=28875.4) and Robot (*M*=29209, *SD*=25298.58) conditions.

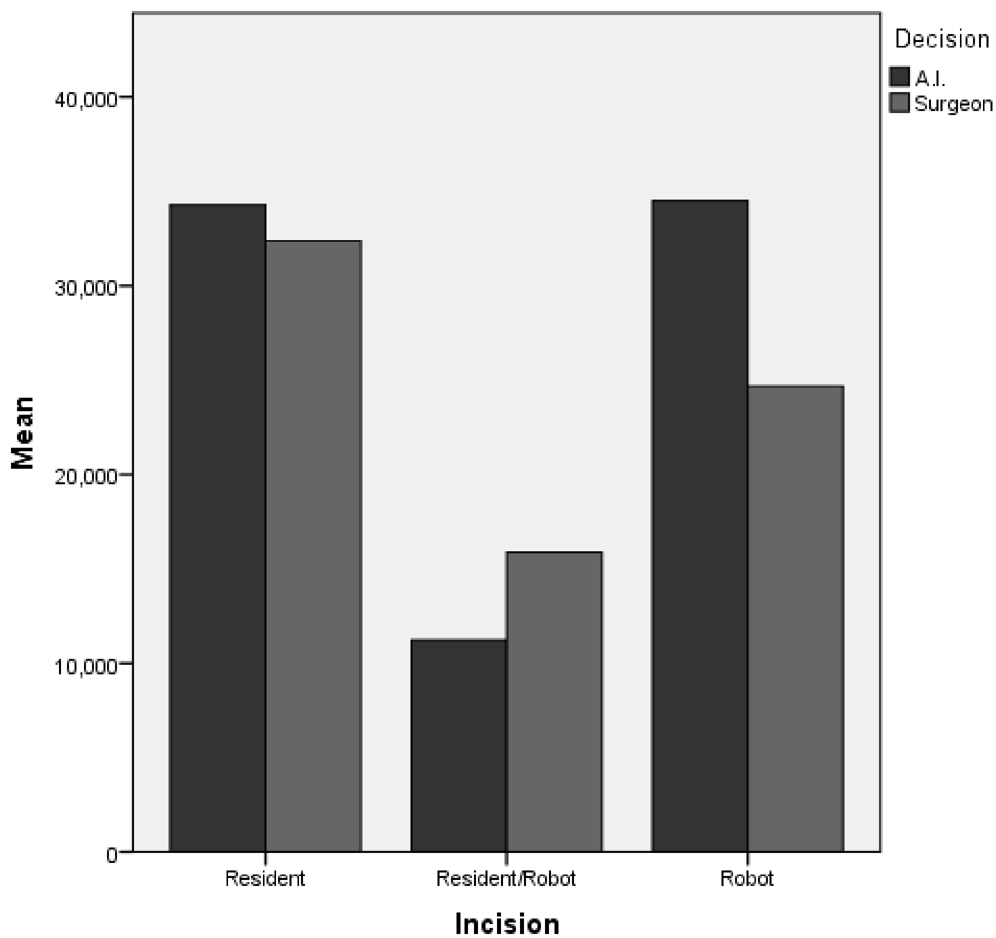


**Figure 7 Graphic of Results for "How much should the patient be compensated for the resulting medical expenses and general damages?"**

# Discussion

The hypothesis was that there existed some underlying effect, in which differences in the type of entity which is in control of a task and the type of entity which is performing a task will lead to different attributions of blame from the perspective of a human observer. This failed to emerge for a large chunk of results. This may be due to the fact that participants were not willing to blame anyone, as all results which asked participants to attribute blame to an entity were below the mid-point of blaming the entity for the post-operative complications. If any significant differences between experimental conditions became apparent, they only served to push the mean lower. This may be due to the fact that participants simply did not understand the vignette or that there were so many entities that it was really no one group's fault. This, along with many of the areas in which statistically significant results did emerge suggests that there may be a diffusion of responsibility. The more entities which are salient in the participant's mind, the less responsibility each holds. This will be discussed in much greater detail later on.

Interestingly, there was a main effect for Decision when participants were asked how likely the Intake Nurse was at fault for post-operative complications. An initial intuition may suggest that participants possibly thought the nurse was charged with supervising the diagnostic A.I., and the Nurse for issues stemming from the A.I. or that participants simply wanted to blame a human. However, the pattern of results suggest something else, as participants in the A.I. condition actually thought it was less likely that the Nurse was at fault than did the participants in the Surgeon condition. A second intuition may suggest that participants thought the nurse felt some sort of unease in correcting a surgeon's diagnosis inherent to the existing social structure, and thus they blame her for not speaking up. And yet, a third intuition may suggest a Type I error due to $\alpha$-inflation, as there were 15 separate ANOVAs and a Kurskal-Wallis that were performed. With 16 statistical tests and statistical significance being set at $\alpha = .05$, there is an approximately 44% chance of at least one incorrect rejection of the null. With the p-value for the main-effect on the two-way ANOVA being .046, this may be the most parsimonious conclusion to be drawn.

However, assuming this was not a Type I error, there still exists a theoretical direction with which this may be taken. Looking at the question set which explored which entity was at fault, there were only two questions which assessed entities as individuals: the patient and the nurse. Due to the fact that the patient is so intertwined with the overarching scenario, those results come with their own confounds. Thus, the nurse is the only entity who may be seen as both equally responsible as any single individual and the only entity whose responsibility was explored, as a question, on the individual level. This result, as well as the results regarding monetary compensation which will be discussed later, may imply a sort of diffusion of responsibility. The more entities that are salient to the participant, the less responsibility each one has.

There were many sections of the vignette which, on further inspection, may inform the results for the patient. There were a handful of lines, such as "…as per the patient's request." and "…the patient decided to undergo…" which may have, unfortunately, implied to the participant that the patient had more choice than was originally intended. Because they possibly saw that the patient had more choice than other entities, they may have perceived the patient to be more agentic, and thus altered their pattern of reporting.

In general, the results took a semi-consistent pattern. Whenever a significant main effect was found for Incision, the means were clustered such that the participants in the Robot condition and Resident condition rated questions similarly, while the participants in the Resident/Robot condition generally rated questions more positively than the other two groups. When participants were asked if they believed the post-operative complications were due to an error, participants in the Robot and Resident replied similarly, believing that an error was somewhere between "less likely than not" and "somewhat likely." However, participants in the Resident/Robot condition answered that the probability of an error was "less likely than not."

When participants were asked how large of a compensation the patient deserved, using a Likert scale, there was no main effect for Incision, but there was one for Decision. Participants in the A.I. condition believe that "the patient deserves a moderately sized compensation." while participants in the Surgeon condition believed that "the patient deserves a small compensation." This result is especially interesting in context of the next question participants were asked: to actually assign a monetary amount for the compensation.

In that question, there was no main effect for Decision like in the previous question. Instead, there was a main effect for Incision. Participants in the Resident condition and the Robot condition had the exact same median compensation size, which was exactly 2.5 times that of the Resident/Robot condition. Looking at the means, this pattern is still relatively consistent: Participants in the Resident/Robot condition believed the patient should be compensated less than the participants in both the Resident and Robot conditions. However, this time, the compensation for the Robot condition was approximately twice as large while the Resident condition was still 2.5 times as large as the Resident/Robot condition. Looking at the Likert scale question, it is possible that the Decision manipulation affected what participants consider, in this context, a moderate monetary amount and small monetary amounts, but the actual amounts they believed the participant deserved stayed constant, relative to the Incision condition. Similarly, the Incision condition may not have affected if the participants believed the patient deserved a moderate or small compensation. It simply redefined what those categories meant to the participants. It is also possible that participants did not know where the compensation was coming from, and specifying this may result in different effects. Additionally, specifying who is paying before asking which entity is at fault will likely prime participants to respond differently and it may be of interest to see how the experimental manipulations interact with the new at-fault entity once participants have something to actually blame.

These results possibly speak to a similar concept that the Intake Nurse question did: a diffusion of responsibility. It is hopefully more than just coincidence that, when placed in a condition with two Incision entities, the robot and resident in the Resident/Robot condition, participants gave the patient almost exactly half as much compensation as they did in either the Robot or Resident conditions. This may imply that participants in the Resident/Robot condition have a mental representation of two separate entities, which would help to support the thought that the participants viewed the resident and robot in collaboration, rather than the robot simply as a tool of the resident. This leads to a possible future direction of research, as it may be of interest to explore definitively whether the participant viewed the robot and resident in the Resident/Robot condition as two separate entities. Additionally, it may also be useful to include a cooperative element within the Decision manipulation, similar to the Resident/Robot condition inside the Incision manipulation.

The current study is limited in a few areas, and thus certain interpretations and results may fail to transfer in subsequent experiments. The largest hole in the current study involves participation. A power analysis performed prior to data collection suggested that approximately 120 participants would be required, while the current study only collected 104 quality participants. There exists the possibility of participants not being as engaged as necessary with the study, as it was taken through an online medium. This may explain the large rate of participants who failed to properly respond to survey questions or failed to finish the survey. Future studies of this nature may find it useful to have participants come in and take the survey in a face to face setting in the hopes of avoiding similar issues.

Another issue may be that the scenario outlined in the vignette may have attempted to maintain too much ecological validity which came at the expense of making it too difficult to read, and thus participants simply did not understand the scenario or that the signal to noise ratio in the vignette was too high and participants were focusing on the distractor details too much. A follow-up may require a different population, such as medical students.

With Florida becoming the second state in 2012 to allow for driverless cars (Valdes, 2012), the study of how other drivers perceive these cars will be essential in the study of how those drivers will react when near one of these cars, when they have a collision with these cars, and when these cars are inevitably in a courtroom scenario. The current study transfers very well to these driverless cars, as the "Incision" manipulation parallels that of a driver: A human can drive a car, an A.I. can drive a car, and in some of these cars, the human can choose when they control the driverless car (National Highway Traffic Safety Administration, 2013). While the "Decision" manipulation is not as easy to transfer, it may still be of interest to explore how the route traveled was generated, whether the driver or some program made the decision. The scenario may end in a manner similar to "post-operative complications," with the trip ending in the car traveling to the wrong destination, and the driver becoming lost or the trip may simply end in a collision.

Overall, the findings of the study performed parallel what you find in the anthropomorphism, human-computer interaction, and human-robot interaction literature: Generally, if you take any social science finding which describes the interaction between two humans, and substitute a non-human agent for one of the humans, the general finding will still apply (Nass, Steuer, & Tauber, 1994). However, there are still some results, both in the study performed herein and the overall literature, which deviate from this principle. This is where further study is warranted. What common elements exist within these scenarios which cause them to deviate from the expected results?

# Appendix

## IRB Outcome Letters

University of Central Florida Institutional Review Board
Office of Research & Commercialization
12201 Research Parkway, Suite 501
Orlando, Florida 32826-3246
Telephone: 407-823-2901 or 407-882-2276
www.research.ucf.edu/compliance/irb.html

### Approval of Exempt Human Research

From: **UCF Institutional Review Board #1**
**FWA00000351, IRB00001138**

To: **Valerie K. Sims** and Co-PI: **Federico Scholcover**

Date: **May 29, 2013**

Dear Researcher:

On 5/29/2013, the IRB approved the following activity as human participant research that is exempt from regulation:

| | |
|---|---|
| Type of Review: | Exempt Determination |
| Project Title: | Observer Perceptions of Agency of Nonliving Beings |
| Investigator: | Valerie K Sims |
| IRB Number: | SBE-13-09402 |
| Funding Agency: | |
| Grant Title: | |
| Research ID: | N/A |

This determination applies only to the activities described in the IRB submission and does not apply should any changes be made. If changes are made and there are questions about whether these changes affect the exempt status of the human research, please contact the IRB. When you have completed your research, please submit a Study Closure request in iRIS so that IRB records will be accurate.

In the conduct of this research, you are responsible to follow the requirements of the Investigator Manual.

On behalf of Sophia Dziegielewski, Ph.D., L.C.S.W., UCF IRB Chair, this letter is signed by:

Signature applied by Joanne Muratori on 05/29/2013 09:51:25 AM EDT

Joanne Muratori

IRB Coordinator

Page 1 of 1

University of Central Florida

University of Central Florida Institutional Review Board
Office of Research & Commercialization
12201 Research Parkway, Suite 501
Orlando, Florida 32826-3246
Telephone: 407-823-2901 or 407-882-2276
www.research.ucf.edu/compliance/irb.html

## Approval of Exempt Human Research

From: **UCF Institutional Review Board #1**
**FWA00000351, IRB00001138**

To: **Valerie K. Sims** and Co-PI: **Federico Scholcover**

Date: **March 05, 2014**

Dear Researcher:

On 3/5/2014, the IRB approved the following minor modification to human participant research that is exempt from regulation:

| | |
|---|---|
| Type of Review: | Exempt Determination |
| Modification Type: | In addition to online setting, study participants can take part in the study in a face-to-face setting in the UCF Psychology Building. Participants will sit in front of a computer, which will link them to the consent document and they will be directed to follow screen prompts. |
| Project Title: | Observer Perceptions of Agency of Nonliving Beings |
| Investigator: | Valerie K Sims |
| IRB Number: | SBE-13-09402 |
| Funding Agency: | |
| Grant Title: | |
| Research ID: | N/A |

This determination applies only to the activities described in the IRB submission and does not apply should any changes be made. If changes are made and there are questions about whether these changes affect the exempt status of the human research, please contact the IRB. When you have completed your research, please submit a Study Closure request in iRIS so that IRB records will be accurate.

In the conduct of this research, you are responsible to follow the requirements of the Investigator Manual.

On behalf of Sophia Dziegielewski, Ph.D., L.C.S.W., UCF IRB Chair, this letter is signed by:

Signature applied by Joanne Muratori on 03/05/2014 01:02:02 PM EST

Joanne Muratori

IRB Coordinator

Page 1 of 1

## Vignettes

*Resident/Surgeon*

Please read the following brief excerpt of a transcript of a medical procedure carefully.
You ***WILL*** be asked questions about what you read:

**Key Phrases:**
*Hematuria*: Blood in urine
*Dysuria*: Painful Urination
*PSA*: Prostate-Specific Antigen
*Adenocarcinoma*: A type of cancer
*Retropubic Prostatectomy*: A surgical procedure requiring anesthesia
Asimov Corporation: A corporation specializing in the development and manufacturing of high-tech surgical tools.

**Transcript:**

**Brief History of the Patient:** The patient is a 45-year-old male with a history of urological issues, including a previous hematuria caused by right renal stones. He reported, to the intake nurse, a return of the hematuria, but with incidences of dysuria and acute pain in his pelvic bone. Blood tests returned with an abnormally high PSA level. The patient was given a preoperative diagnosis of adenocarcinoma of the prostate by the attending surgeon. Upon consideration of various treatment modalities and their associated preoperative prognoses, the patient decided to undergo a retropubic prostatectomy, in order to remove the adenocarcinoma.

**Procedure:** The patient was taken into the operating room, and successfully given a general anesthetic. As per the patient's request, the surgery was performed by a 5th year surgical resident under the supervision of the attending surgeon. The surgery was performed to standard.

**Postoperative Status:** Post-surgical blood tests have indicated a return to normal PSA levels, as well as no signs of the adenocarcinoma. However, due to the nature of the procedure, the patient has experienced postoperative complications.

*Resident and Robot/Surgeon*

Please read the following brief excerpt of a transcript of a medical procedure carefully.
You ***WILL*** be asked questions about what you read:

**Key Phrases:**
*Hematuria*: Blood in urine
*Dysuria*: Painful Urination
*PSA*: Prostate-Specific Antigen
*Adenocarcinoma:* A type of cancer
*Retropubic Prostatectomy*: A surgical procedure requiring anesthesia
Asimov Surgical System: A robotic system with five surgical arms, each capable of performing different surgical tasks, including incisions and endoscopies.
Asimov Corporation: A corporation specializing in the development and manufacturing of high-tech surgical tools.

**Transcript:**

**Brief History of the Patient:** The patient is a 45-year-old male with a history of urological issues, including a previous hematuria caused by right renal stones. He reported, to the intake nurse, a return of the hematuria, but with incidences of dysuria and acute pain in his pelvic bone. Blood tests returned with an abnormally high PSA level. The patient was given a preoperative diagnosis of adenocarcinoma of the prostate by the attending surgeon. Upon consideration of various treatment modalities and their associated preoperative prognoses, the patient decided to undergo a retropubic prostatectomy, in order to remove the adenocarcinoma.

**Procedure**: The patient was taken into the operating room, and successfully given a general anesthetic. As per the patient's request, the surgery was performed by a 5th year surgical resident controlling an Asimov Surgical System under the supervision of the attending surgeon. The surgery was performed to standard.

**Postoperative Status:** Post-surgical blood tests have indicated a return to normal PSA levels, as well as no signs of the adenocarcinoma. However, due to the nature of the procedure, the patient has experienced postoperative complications.

*Robot/Surgeon*

Please read the following brief excerpt of a transcript of a medical procedure carefully.
You ***WILL*** be asked questions about what you read:

**Key Phrases:**
*Hematuria*: Blood in urine
*Dysuria*: Painful Urination
*PSA*: Prostate-Specific Antigen
*Adenocarcinoma:* A type of cancer
*Retropubic Prostatectomy*: A surgical procedure requiring anesthesia
Asimov Surgical System: A robotic system with five surgical arms, each capable of performing different surgical tasks, including incisions and endoscopies.
Asimov Corporation: A corporation specializing in the development and manufacturing of high-tech surgical tools.

**Transcript:**

**Brief History of the Patient:** The patient is a 45-year-old male with a history of urological issues, including a previous hematuria caused by right renal stones. He reported, to the intake nurse, a return of the hematuria, but with incidences of dysuria and acute pain in his pelvic bone. Blood tests returned with an abnormally high PSA level. The patient was given a preoperative diagnosis of adenocarcinoma of the prostate by the attending surgeon. Upon consideration of various treatment modalities and their associated preoperative prognoses, the patient decided to undergo a retropubic prostatectomy, in order to remove the adenocarcinoma.

**Procedure**: The patient was taken into the operating room, and successfully given a general anesthetic. As per the patient’s request, the surgery was performed by a fully automated Asimov Surgical System under the supervision of the attending surgeon. The surgery was performed to standard.

**Postoperative Status:** Post-surgical blood tests have indicated a return to normal PSA levels, as well as no signs of the adenocarcinoma. However, due to the nature of the procedure, the patient has experienced postoperative complications.

*Resident/A.I.*

Please read the following brief excerpt of a transcript of a medical procedure carefully.
You ***WILL*** be asked questions about what you read:

**Key Phrases:**
*Hematuria*: Blood in urine
*Dysuria*: Painful Urination
*PSA*: Prostate-Specific Antigen
*Adenocarcinoma:* A type of cancer
*Retropubic Prostatectomy*: A surgical procedure requiring anesthesia
Asimov Surgical A.I.: An artificial intelligence developed by the Asimov Corporation, capable of performing analytical tasks, such as diagnostics and surgical supervision.
Asimov Corporation: A corporation specializing in the development and manufacturing of high-tech surgical tools.

**Transcript:**

**Brief History of the Patient:** The patient is a 45-year-old male with a history of urological issues, including a previous hematuria caused by right renal stones. He reported, to the intake nurse, a return of the hematuria, but with incidences of dysuria and acute pain in his pelvic bone. Blood tests returned with an abnormally high PSA level. The patient was given a preoperative diagnosis of adenocarcinoma of the prostate by the Asimov Surgical A.I. Upon consideration of various treatment modalities and their associated preoperative prognoses, the patient decided to undergo a retropubic prostatectomy, in order to remove the adenocarcinoma.

**Procedure**: The patient was taken into the operating room, and successfully given a general anesthetic. As per the patient's request, the surgery was performed by a 5th year surgical resident under the supervision of the Asimov Surgical A.I. The surgery was performed to standard.

**Postoperative Status:** Post-surgical blood tests have indicated a return to normal PSA levels, as well as no signs of the adenocarcinoma. However, due to the nature of the procedure, the patient has experienced postoperative complications.

*Resident and Robot/A.I.*

Please read the following brief excerpt of a transcript of a medical procedure carefully.
You ***WILL*** be asked questions about what you read:

**Key Phrases:**
*Hematuria*: Blood in urine
*Dysuria*: Painful Urination
*PSA*: Prostate-Specific Antigen
*Adenocarcinoma:* A type of cancer
*Retropubic Prostatectomy*: A surgical procedure requiring anesthesia
Asimov Surgical System: A robotic system with five surgical arms, each capable of performing different surgical tasks, including incisions and endoscopies.
Asimov Surgical A.I.: An artificial intelligence developed by the Asimov Corporation, capable of performing analytical tasks, such as diagnostics and surgical supervision.
Asimov Corporation: A corporation specializing in the development and manufacturing of high-tech surgical tools.

**Transcript:**

**Brief History of the Patient:** The patient is a 45-year-old male with a history of urological issues, including a previous hematuria caused by right renal stones. He reported, to the intake nurse, a return of the hematuria, but with incidences of dysuria and acute pain in his pelvic bone. Blood tests returned with an abnormally high PSA level. The patient was given a preoperative diagnosis of adenocarcinoma of the prostate by the Asimov Surgical A.I. Upon consideration of various treatment modalities and their associated preoperative prognoses, the patient decided to undergo a retropubic prostatectomy, in order to remove the adenocarcinoma.

**Procedure**: The patient was taken into the operating room, and successfully given a general anesthetic. As per the patient's request, the surgery was performed by a 5th year surgical resident controlling an Asimov Surgical System under the supervision of the Asimov Surgical A.I. The surgery was performed to standard.

**Postoperative Status:** Post-surgical blood tests have indicated a return to normal PSA levels, as well as no signs of the adenocarcinoma. However, due to the nature of the procedure, the patient has experienced postoperative complications.

*Robot/A.I.*

Please read the following brief excerpt of a transcript of a medical procedure carefully.
You ***WILL*** be asked questions about what you read:

**Key Phrases:**
*Hematuria*: Blood in urine
*Dysuria*: Painful Urination
*PSA*: Prostate-Specific Antigen
*Adenocarcinoma:* A type of cancer
*Retropubic Prostatectomy*: A surgical procedure requiring anesthesia
Asimov Surgical System: A robotic system with five surgical arms, each capable of performing different surgical tasks, including incisions and endoscopies.
Asimov Surgical A.I.: An artificial intelligence developed by the Asimov Corporation, capable of performing analytical tasks, such as diagnostics and surgical supervision.
Asimov Corporation: A corporation specializing in the development and manufacturing of high-tech surgical tools.

**Transcript:**

**Brief History of the Patient:** The patient is a 45-year-old male with a history of urological issues, including a previous hematuria caused by right renal stones. He reported, to the intake nurse, a return of the hematuria, but with incidences of dysuria and acute pain in his pelvic bone. Blood tests returned with an abnormally high PSA level. The patient was given a preoperative diagnosis of adenocarcinoma of the prostate by the Asimov Surgical A.I. Upon consideration of various treatment modalities and their associated preoperative prognoses, the patient decided to undergo a retropubic prostatectomy, in order to remove the adenocarcinoma.

**Procedure**: The patient was taken into the operating room, and successfully given a general anesthetic. As per the patient's request, the surgery was performed by a fully automated Asimov Surgical System under the supervision of the Asimov Surgical A.I. The surgery was performed to standard.

**Postoperative Status:** Post-surgical blood tests have indicated a return to normal PSA levels, as well as no signs of the adenocarcinoma. However, due to the nature of the procedure, the patient has experienced postoperative complications.

**Vignette Questions**[1]

The patient had a history of _______

- ❍ Urological Issues
- ❍ Schizotypal Personality Disorder
- ❍ Irritable Bowel Syndrome

The goal of the procedure was ______

- ❍ Removal of the adenocarcinoma
- ❍ Lobotomy
- ❍ Exploratory Surgery

The patient experienced post-operative complications

- ❍ True
- ❍ False

[1] Please note, participants were not presented with the entire set of questions at once. Thus, a new page indicates that participants were presented with a new set of questions. Participants were unable to return any of the previous set of questions.

Imagine that you're now a juror for a trial regarding the very same medical procedure you just read about. How would you answer the following questions?

How likely is it that the post-operative complications were due to an error?

- Extremely Unlikely
- Less Likely than Not
- Somewhat Likely
- More Likely than Not
- Extremely Likely

How likely is it that the post-operative complications were due to random chance?

- Extremely Unlikely
- Less Likely than Not
- Somewhat Likely
- More Likely than Not
- Extremely Likely

Imagine that you're now a juror for a trial regarding the very same medical procedure you just read about. How would you answer the following questions?

How likely is it that there was an error in the diagnosis prior to the surgery being performed?

- Extremely Unlikely
- Less Likely than Not
- Somewhat Likely
- More Likely than Not
- Extremely Likely

How likely is it that the correct surgical procedure was performed?

- Extremely Unlikely
- Less Likely than Not
- Somewhat Likely
- More Likely than Not
- Extremely Likely

How likely is it that an error occurred after the surgery?

- Extremely Unlikely
- Less Likely than Not
- Somewhat Likely
- More Likely than Not
- Extremely Likely

Imagine that you're now a juror for a trial regarding the very same medical procedure you just read about. How would you answer the following questions?

How likely is it that a mental error occurred during surgery?

- Extremely Unlikely
- Less Likely than Not
- Somewhat Likely
- More Likely than Not
- Extremely Likely

How likely is it that a physical error occurred during surgery?

- Extremely Unlikely
- Less Likely than Not
- Somewhat Likely
- More Likely than Not
- Extremely Likely

Imagine that you're now a juror for a trial regarding the very same medical procedure you just read about. How would you answer the following questions?

How likely are you to find the hospital at fault for the post-operative complications?

- Extremely Unlikely
- Less Likely than Not
- Somewhat Likely
- More Likely than Not
- Extremely Likely

How likely are you to find Asimov Corporation at fault for the post-operative complications?

- Extremely Unlikely
- Less Likely than Not
- Somewhat Likely
- More Likely than Not
- Extremely Likely

Imagine that you're now a juror for a trial regarding the very same medical procedure you just read about. How would you answer the following questions?

How likely are you to find the diagnostic team, as a whole, at fault for the post-operative complications?

- ❍ Extremely Unlikely
- ❍ Less Likely than Not
- ❍ Somewhat Likely
- ❍ More Likely than Not
- ❍ Extremely Likely

How likely are you to find the surgical team, as a whole, at fault for the post-operative complications?

- ❍ Extremely Unlikely
- ❍ Less Likely than Not
- ❍ Somewhat Likely
- ❍ More Likely than Not
- ❍ Extremely Likely

How likely are you to find the patient at fault for the post-operative complications?

- ❍ Extremely Unlikely
- ❍ Less Likely than Not
- ❍ Somewhat Likely
- ❍ More Likely than Not
- ❍ Extremely Likely

How likely are you to find the intake nurse at fault for the post-operative complications?[1]

- ❍ Extremely Unlikely
- ❍ Less Likely than Not
- ❍ Somewhat Likely
- ❍ More Likely than Not
- ❍ Extremely Likely

[1] Question presented with the questions on the previous page.

Imagine that you're now a juror for a trial regarding the very same medical procedure you just read about. How would you answer the following questions?

How much compensation does the patient deserve for the post-operative complications?

- ❍ The patient deserves no compensation.
- ❍ The patient deserves a small compensation.
- ❍ The patient deserves a moderately sized compensation.
- ❍ The patient deserves a large compensation.
- ❍ The patient deserves as large of a compensation as he can get.

Assume the patient will be receiving compensation for the resulting medical expenses. How much should the patient be compensated for the resulting medical expenses and general damages? Please enter a number between $0 and $100,000.

## Works Cited


Ballantyne, G. H. (2002). Robotic surgery, telerobotic surgery, telepresence, and telementoring. *Surgical Endoscopy and Other Interventional Techniques*, *16*(10), 1389-1402.

Breazeal, C. (2004). Social interactions in HRI: the robot view. Systems, Man, and Cybernetics, Part C: Applications and Reviews, *IEEE Transactions on*, *34*(2), 181-186.

Donnellan, M.B., Oswald, F.L., Baird, B.M., & Lucas, R.E. (2006). The mini-IPIP scales: Tiny-yet-effective measures of the Big Five factors of personality. *Psychological Assessment, 18*, 192-203.

Duffy, B. R. (2003). Anthropomorphism and the social robot. *Robotics and autonomous systems*, *42*(3), 177-190.

Epley, N., Waytz, A., & Cacioppo, J. T. (2007). On seeing human: a three-factor theory of anthropomorphism. *Psychological review*, *114*(4), 864.

Eyssel, F., & Kuchenbrandt, D. (2011). Social categorization of social robots: Anthropomorphism as a function of robot group membership. *British Journal of Social Psychology*, *51*(4), 724-731.

Eyssel, F., Kuchenbrandt, D., & Bobinger, S. (2011, March). Effects of anticipated human-robot interaction and predictability of robot behavior on perceptions of anthropomorphism. *Proceedings of the 6th international conference on Human-robot interaction*, 61-68.

Gherhardus, D. (2003). Robot assisted surgery: The future is here. *Journal of Healthcare Management*, *48*(4), 242-251.

Gong, L. (2008). How social is social responses to computers? The function of the degree of anthropomorphism in computer representations. *Computers in Human Behavior*, *24*(4), 1494-1509.

Hurwitz, M. (Producer) (2005). The immaculate election [Television series episode]. In Hurwitz, M. (Executive Producer), Arrested Development. Los Angeles, CA: Fox Broadcasting Company.

Kim, T., & Hinds, P. (2006, September). Who should I blame? Effects of autonomy and transparency on attributions in human-robot interaction. *Robot and Human Interactive Communication, 2006. The 15th IEEE International Symposium*, 80-85.

Lanfranco, A. R., Castellanos, A. E., Desai, J. P., & Meyers, W. C. (2004). Robotic surgery: a current perspective. *Annals of surgery, 239(*1), 14.

MTSamples.com. (n.d.). Retrieved March 9, 2014 from the MTSamples: http://www.mtsamples.com/

Nass, C., Steuer, J., & Tauber, E. R. (1994, April). Computers are social actors. *Proceedings of the SIGCHI conference on Human factors in computing systems,* 72-78.

National Highway Traffic Safety Administration. (2013). U.S. Department of Transportation Releases Policy on Automated Vehicle Development. [Press release]. Retrieved from http://www.nhtsa.gov/About+NHTSA/Press+Releases/U.S.+Department+of+Transportation+Releases+Policy+on+Automated+Vehicle+Development

Nomura, T., Kanda, T., & Suzuki, T. (2006). Experimental investigation into influence of negative attitudes toward robots on human–robot interaction. *AI & Society, 40,* 138-150. doi: 10.1007/s00146-005-0012-7

Nowak, K. L., & Biocca, F. (2003). The effect of the agency and anthropomorphism on users' sense of telepresence, copresence, and social presence in virtual environments. *Presence: Teleoperators and Virtual Environments, 12*(5), 481-494.

Principe, C. P., & Langlois, J. H. (2013). Children and adults use attractiveness as a social cue in real people and avatars. *Journal of Experimental Child Psychology,* http://dx.doi.org/10.1016/j.jecp.2012.12.002

Rotter, J. B. (1966). Generalized expectancies for internal versus external control of reinforcements. *Psychological Monographs: General and Applied*, *80*(1), 1-28.

Syrdal, D. S., Koay, K. L., Walters, M. L., & Dautenhahn, K. (2007, August). A personalized robot companion? The role of individual differences on spatial preferences in HRI scenarios. *Robot and Human interactive Communication, 2007. The 16th IEEE International Symposium*, 1143-1148.

Valdes, A. M. (2012, May 07). Florida embraces self-driving cars, as engineers and lawmakers prepare for the new technology. *WPTV NewsChannel 5*. Retrieved from http://www.wptv.com/news/state/florida-embraces-self-driving-cars-as-engineers-and-lawmakers-prepare-for-the-new-technology

Yee, Nick, Jeremy N. Bailenson, and Kathryn Rickertsen. (2007). "A meta-analysis of the impact of the inclusion and realism of human-like faces on user experiences in interfaces." *Proceedings of the SIGCHI conference on Human factors in computing systems*.