Robot, Rabbit, or Red Herring? Societal Acceptance as a Function of Classification Ease

Laurel D. Riek, *Member, IEEE* and Peter Robinson, *Member, IEEE*

***Abstract*—For people to accept robotic agents socially it is necessary for the robots to be easily classifiable. We propose three such kinds of classification for users: Type Classification (what is it?), Role Classification (how should I interact with it?), and Behavioral Classification (does it behave in concordance with its type and role?). As HRI researchers we can design experiments that measure length of time until people classify a given robot's type and role, and a robot's behavior in terms of the number of classification violations it commits. These measurements can then be used to calculate an Ease of Classification score. This score could be used as a basis for comparison between different user groups, different physical interaction spaces, and even different robots. Further, it can help provide insight into the likelihood a given robot will be socially accepted.**

I. Introduction

THERE are a variety of tools available to Human-Robot Interaction (HRI) researchers seeking to assess aspects of the societal acceptance of robots. Some successful techniques described in the literature include: ethnographic observation [1], system response-time analysis [2], common ground analysis [3], embodiment measurement [4], perceived enjoyment analysis [5], comfort level analysis [6], interaction profile analysis [7], and others [8].

We suggest a new addition to this toolset: Classification Ease. For people to accept robots in social contexts it is important that the robots be *easily classifiable*, that is – end users should be able to quickly and easily identify a robot's type, role, and behavioral function. We hypothesize that users will be more apt to feel comfortable around a robot that is easily classifiable, and thus will be more accepting of it.

Classification Ease is consistent with one of the core ideas in Human-Centered Design: that technology acceptance is directly related to consistency with users' mental models. In other words, the user should always be able to figure out what to do and know what is going on [9] with regards to interacting with the robot. It is also motivated by the Cognitive Dimensions Framework [10]. This is a "broad-brush" evaluation technique described in the Human-Computer Interaction (HCI) literature as a means for designers who are not HCI experts to make evaluations that are quick but still very useful [11]. For our purpose, Type Classification corresponds with the Perceptual Mapping Dimension (what the robot's physical appearance conveys), Role Classification with the Role-expressiveness Dimension (what role the robot presents), and Behavioral Classification with the Closeness of Mapping Dimension (how the robot's behavior maps to its type and role).

HRI researchers can design very straightforward experiments to measure how long it takes a user to classify a robot's type and role, as well as the robot's behavior in terms of the number of ways in which it violates the user's classifications. Researchers can then use these measurements to calculate an Ease of Classification (EOC) score.

Using an EOC score as a means for measuring societal acceptance of robots has several advantages. First, it can be used as a basis for comparison between different user groups, different physical interaction spaces, and even different robots. Second, in contrast to some of the aforementioned HRI assessment methods it is relatively quick and easy to measure. Last, it is designed to be a flexible metric that can accommodate the needs of different user groups and different user types.

Fig. 1. The Robotceptionist. Note the flowers, business cards, and memorabilia surrounding the desk, as well as the office attire the robot is dressed in. Photograph © 2005 by IEEE.

Manuscript received May 30, 2008. This work was supported in part by the QUALCOMM Research Studentship in Computing.

L. D. Riek is with the Cambridge University Computer Laboratory, Cambridge, CB3 0FD UK. (Email: Laurel.Riek@cl.cam.ac.uk)

P. Robinson is with the Cambridge University Computer Laboratory, Cambridge, CB3 0FD UK. (Email: Peter.Robinson@cl.cam.ac.uk)

II. Kinds of Classification

A. Type Classification

Upon first encountering a robot within its intended physical space, it should be immediately apparent to users what general purpose the robot is intended to serve, i.e. its type. The robot's physical appearance, movement, gait, speech, gesture, gaze, or stature can reflect this purpose. For example, when one encounters Tank the Roboceptionist [12] (Fig. 1) at Carnegie Mellon it is very easy to classify its type as 'receptionist'. Tank is located near the entrance to a building, is stationary, is situated at human height inside a wooden booth, and is surrounded by objects that one would usually encounter in an office environment (business cards, vase of flowers, etc). Tank's interaction style closely mirrors one of a friendly human receptionist in terms of polite speech, raised eyebrows, and a desire to answer questions. This robot is unlikely to be mistaken for anything other than what it is because its design and physical placement clearly reflect its type and purpose.

In contrast, the fictional talking car KITT depicted on the television show *Knight Rider* often presented a very misleading type. When first glancing at KITT people assumed it was merely a car, not a socially interactive robot. The sole visual cue to its function was an LED display of lights inside the vehicle that only illuminated when it was speaking. Often KITT would remain silent and then suddenly speak, which startled people. This is because they had already classified its type as a non-sentient vehicle based upon the visual cues afforded to them.

To help ensure a robot presents itself in a way that its type is easily and quickly classifiable, it may be helpful for robot designers to employ at least rudimentary contextual design [13] techniques. The more this process is employed the less guesswork will ultimately be required by end users, thus ensuring a greater likelihood of acceptance. Interestingly, alterations informed by contextual design needn't be elaborate; otherwise physically unremarkable robots tele-operated in hospitals by remotely located physicians are often "dressed up" in white lab coats by hospital staff to allow patients to quickly identify their function [14]. This ease of identification through physical means has surely helped the patients accept the presence of robots.

Indeed, people who interact with personal robots in the home will often dress them in costumes [15], perhaps as a means to help other family members and visitors to the home readily classify the robot as non-threatening.

B. Role Classification

Role Classification refers to the kind of interactive relationship a user might expect from a robot. Goodrich and Schultz suggest the general roles interactive robots serve may be peer, assistant, or slave, and that the robot may report to another robot or human, or be fully independent [8]. Scholtz et. al. propose a taxonomy of roles that robots can be: Supervisor, Operator, Mechanic, Peer, and Bystander [16]; robots may also assume roles such as Mentor or Information Consumer [8].

Some examples of robots that present easily classifiable roles include the iRobot Roomba™ vacuum cleaner (slave), the Sage museum guide robot [17] (mentor), and the PARO robotic seal [18] (peer).

The role(s) a robot will adopt during interaction with a user should always be made explicit and should be immediately apparent. Furthermore, a robot should remain consistent with its advertised role. A comforting companion robot in a nursing home serving in a peer role should not suddenly adopt a mentor role and start lecturing residents about their health habits.

The correct assignment of a robot's role depends entirely upon the context in which it will be used and the population of users it is intended to interact with. Again, robot designers will benefit from closely studying the environment in which they wish to deploy their robot before embarking on its creation.

Fig. 2. The PARO robotic seal robot. This robot is embodied within a stuffed animal, and acts as one might expect an animate toy to behave. It has generally been well accepted as a companion robot. Photograph © 2005 ACM.

C. Behavioral Classification

A well-designed robot will present its behavioral function clearly to a user and behave in concordance with its Type and Role Classifications. For example, PARO the robotic seal (Fig. 2) [18] and the Huggable robot [19] present themselves embodied within stuffed animals, and behave as one might expect an animate toy or pet to act. Their behaviors do not violate their Type or Role Classifications, and thus are generally well accepted.

In contrast, there have been several robots described in the literature that elicit a sense of unease among users due to a mismatch between their form, interactivity, and motion quality [20]. This parallels the uncanny valley effect reported by Mori [21]. These are robots that appear very human-like at first glance but then act in very un-humanlike ways. Such robots can be said to strongly violate their Type and Role Classifications, thereby preventing users from easily classifying them. Thus it is understandable that these robots seem not to be well accepted by society.

This is not to say robots ought never to surprise the users interacting with them. If a museum robot were to break from a lecture about dinosaurs and start telling archeology jokes, it is unlikely any humans would experience unease. So long as a robot follows the social conventions and constraints suggested by its type and role, unexpected behaviors are perfectly reasonable.

III. Ease of Classification Score

We propose the following formula to calculate an Ease of Classification (EOC) score. This score is based upon the time it takes a user to identify a robot's type and role, and the count of behavioral violations the robot commits. A robot with a lower EOC score can be said to be more easily classifiable than a robot with a higher one.

Variables

- t_t : Time until type classification by user (in seconds)
- t_r : Time until role classification by user (in seconds)
- t_i : Total time user observes robot (in seconds)
- b : Number of behavioral violations
- C : A constant to help weight behavioral violations
- E : EOC Score

Assumptions

1) Since b is simply a count of behavioral violations we will assume it to be a non-negative integer:

$$b \geq 0$$

2) A user should be able to classify a robot's type and role within five minutes of interacting with it:

$$t_t, t_r, t_i \leq 300\,s$$

EOC Score Formula

The ease of classification score may be calculated with the following formula:

$$E = \max(t_t,\ t_r) + C \bullet b / t_i$$

Notes

As this formula has yet to be verified experimentally, it is presently unclear what C ought to be. We expect this constant will vary depending on the users, robots, and tasks being evaluated. A low C value might be appropriate for an entertainment Bystander robot where behavioral violations are less critical, whereas a high C value might be required for a peer Mechanic robot. In other words, C helps to weight the robot within its context.

We have weighted t_t and t_r equally because we expect users to recognize a robot's type and role at roughly the same time. Were a situation to warrant serial identification, the following formula might be more appropriate:

$$E = t_t + \ t_r + C \bullet b / t_i$$

Such situations might be, for instance, a case of delayed introduction to the robot. For example, a user might see a still image of a robot that readily conveys its type through physical appearance but cannot identify its role until an in-person encounter takes place.

IV. Discussion

We have proposed a new addition to the HRI researcher's toolset: Classification Ease. We hypothesize that when users can easily classify a robot's type, role, and behavioral function they are more apt to feel comfortable with it, and thus will be more accepting of it. This notion is consistent with a core idea in Human-Centered Design, that technology acceptance is directly related to consistency with users' mental models. It is also informed by the Cognitive Dimensions framework, mainly: the Perceptual Mapping Dimension (Type Classification), Role-expressiveness Dimension (Role Classification), and Closeness of Mapping Dimension (Behavioral Classification).

HRI researchers can design relatively straightforward experiments that measure how long it takes a user to classify a robot's type and role, as well as the robot's behavior in terms of the number of ways in which it violates the user's classifications. Researchers can then use these measurements to calculate an Ease of Classification (EOC) score.

The EOC score has several advantages. First, it allows researchers a quantifiable metric that can be used as a basis of comparison between different user groups, different physical interaction spaces, and even different robots. For example, one might wish to compare EOC scores of two different types of companion robots used in both a nursing home and an elementary school. Thus an EOC Score can potentially provide a means for normalizing across varied user groups and robots.

Another advantage to the EOC Score is that it should be a relatively easy metric for researchers to calculate. One possible experimental design would start with priming participants on the general types and roles one might expect to see in a robot before interaction takes place. Subjects would then be allowed unstructured interaction with the robot and asked to state when they have determined the robot's type and role. These times would be noted. The subjects can be asked to think aloud [22] during the experiment and the experimenter can note when they express discomfort, increasing the behavioral violation count accordingly.

A few ideas require further investigation. First, the EOC score in its present form is designed to deal with users' "first impressions" of a robot. As a user acclimates to a robot it is likely their views about it will change over time. Second, it is presently unclear how straightforward it will be to adjust the EOC score to fully accommodate different contexts, user bases, and robot types while remaining an evenly normalized metric. Hopefully adjusting C will be a sufficient way to normalize the EOC score, but if not perhaps an additional constant should be included in the formula.

The hypothesized relationship between Classification Ease and the social acceptance of robots will require experimental validation. We expect that additional research in this area will lead to a refinement of the EOC Score formula and to the establishment of a reproducible testing methodology.

ACKNOWLEDGMENT

L.D. Riek thanks Keith Irwin at North Carolina State University for his assistance.

REFERENCES

[1] J. Forlizzi, "How Robotic Products Become Social Products: An Ethnographic Study of Cleaning in the Home," Proceedings of the 2nd ACM/IEEE International Conference on Human Robot Interaction (HRI 2007), pp. 129-136, ACM/IEEE, March, 2007.

[2] T. Shiwa, T. Kanda, M. Imai, H. Ishiguro, and N. Hagita, "How Quickly Should Communication Robots Respond?," Proceedings of the 3rd ACM/IEEE International Conference on Human Robot Interaction (HRI 2008), pp. 153-160, ACM/IEEE, March, 2008.

[3] K. Stubbs, P. Hinds, and D. Wettergreen, "Autonomy and Common Ground in Human-Robot Interaction: A Field Study", *IEEE Intelligent Systems*, volume 22, number 2, pp. 42-50, March, 2007.

[4] K. Dautenhahn, B. Ogden, and T. Quick, "From Embodied to Socially Embedded Agents – Implications for Interaction-Aware Robots," *Cognitive Systems Research*, Volume 3, Number 3, pp. 397-428, Special Issue on Situated and Embodied Cognition, guest-editor: Tom Ziemke, Elsevier. 2002.

[5] M. Heerink, Kröse, B. Wielinga, and V. Evers, "Enjoyment Intention to Use and Actual Use of a Conversational Robot by Elderly People," Proceedings of the 3rd ACM/IEEE International Conference on Human Robot Interaction (HRI 2008), pp. 113-120, ACM/IEEE, March, 2008.

[6] K. Koay, M. Walters, and K. Dautenhahn, "Methodological Issues Using a Comfort Level Device in Human-Robot Interactions", Proceedings of the 14th IEEE International Workshop on Robot and Human Interactive Communication (RO-MAN 2005), pp. 359-364, IEEE Press, August, 2005.

[7] B. Robins, K. Dautenhahn, R. te Boekhorst, and A. Billard, "Robotic Assistants in Therapy and Education of Children with Autism: Can a Small Humanoid Robot Help Encourage Social Interaction Skills?", *Access In the Information Society* (UAIS), Volume 4, Number 4, pp. 2199-2204, Springer-Verlag, 2005.

[8] M. A. Goodrich and A. C. Schultz, "Human-Robot Interaction: A Survey," *Foundations and Trends in Human-Computer Interaction*, 1(3), pp. 203-275, 2007.

[9] D. Norman, *The Psychology of Everyday Things*. Basic Books, New York, 1988. In paperback as *The Design of Everyday Things*. Doubleday, New York, 1990.

[10] T.R.G. Green, "Cognitive Dimenions Of Notations," In A. Sutcliffe and L. Macaulay (Eds.) *People and Computers V*, Cambridge University Press. Cambridge, 1989.

[11] M.D. Dunlop (Ed), "An Introduction to the Cognitive Dimensions Framework," *Proceedings of the Second Mira Workshop (Monselice, Italy)*, University of Glasgow Computing Science Research Report TR-1997-2, November, 1996. http://www.dcs.gla.ac.uk/mira/workshops/padua_procs

[12] R. Gockley, et. al., "Designing Robots for Long-Term Social Interaction," Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2005), pp. 2199 – 2204, IEEE, August, 2005.

[13] H. Beyer and K. Holzblatt, *Contextual Design: Defining Customer-Centered Systems*. San Francisco, CA: Morgan Kaufmann, 1998.

[14] S. Sties. "'R2D2' Develops Bedside Manner," *Shawnee Dispatch: March 30, 2005.*

[15] J. Sung, R. Grinter, H. Christensen, and L. Guo. "Housewives or Technophiles? Understanding Domestic Robot Owners," Proceedings of the 3rd ACM/IEEE International Conference on Human Robot Interaction (HRI 2008), , pp. 129-136, ACM/IEEE, March, 2008.

[16] J. Scholtz, M. Theofanos, and B. Antonishek, "Theory and Evaluation of Human Robot Interactions," Proceedings of the 36th International Conference on Systems Sciences, Hawaii: IEEE, 2002.

[17] T. Willeke, C. Kunz, and I. Nourbakhsh, "The History of the Mobot Museum Robot Series: An Evolutionary Study," In *Proceedings of FLAIRS 2001*. Florida, 2001.

[18] P. Marti, A. Pollini, A. Rullo, and T. Shibata, "Engaging with Artificial Pets", In Proceedings of the 2005 ACM Annual Conference on European Association of Cognitive Ergonomics, ACM, 2005.

[19] W.D. Stiehl, J. Lieberman, C. Breazeal, L. Basel, L. Lalla, and M. Wolf, "The Design of the Huggable: A Therapeutic Robotic Companion for Relational, Affective Touch," Proceedings of AAAI 2005 Fall Symposium on Caring Machines, AAAI, 2005.

[20] C. Ho, K. F. MacDorman, and Z. A. Pramono, "Human Emotion and the Uncanny Valley: a GLM, MDS, and Isomap Analysis of Robot Video Ratings," In Proceedings of the 3rd ACM/IEEE international Conference on Human Robot interaction (HRI '08), ACM, 2008.

[21] M. Mori. "Bukimi no Tami [The Uncanny Valley]," *Energy*, Volume 7, pp. 33-35. 1970.

[22] J. Preece, Y. Rogers, H. Sharp, D. Benyon, S. Holland, and T. Carey, (1994) *Human-Computer Interaction*. Addison-Wesley, Wokingham, UK, 1994. Pp. 622