# Why Should We Imitate Robots?

# (Extended Abstract)

Yasser Mohammad
Assiut University, Egypt
yasserm@aun.edu.eg

Toyoaki Nishida
Kyoto University, Japan
nishida@i.kyoto-u.ac.jp

## ABSTRACT

Previous research in HRI have shown that human's subjective evaluation of robot's abilities affect the way people interact with robots. Given that one of the major challenges in learning from demonstration in robotics is the limited number of training examples that the demonstrator is usually willing to provide, it would be beneficial to design the interaction context in such a way to increase human's subjective evaluation of the robot's imitative skills. We propose back imitation as a way to achieve that goal. This paper reports the results of a preliminary study that was conducted to evaluate the effect of back imitation on human's subjective evaluation of the robot along several dimensions including imitation skill, motion human likeness, interaction quality, humanness and likability.

## Categories and Subject Descriptors

H.4 [**Information Systems Applications**]: Miscellaneous

## General Terms

Experimentation, Human Factors

## Keywords

Human-Robot Interaction, Imitation, back imitation

## 1. INTRODUCTION

A common challenge in learning by demonstration is the need to learn from as few demonstrations as possible [1]. To allow the robot to get more demonstrations, it is beneficial to increase the demonstrator's acceptance and willingness to interact with the robot as much as possible through careful design of the interaction.

The Technology Acceptance Model (TAM) [2] proposes two main variables that affect acceptance: perceived usefulness and perceived ease of use. In the case of a robot learning through imitation, increased perception of imitative skill leads directly to increased performance expectancy (usefulness) and decreased effort expectancy (increased ease of use).

**Appears in:** *Alessio Lomuscio, Paul Scerri, Ana Bazzan, and Michael Huhns (eds.), Proceedings of the 13th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2014), May 5-9, 2014, Paris, France.*


This means that interactions with imitative robots should be designed to increase the perception of robot's imitative skills in order to increase the acceptance of the robot. This paper proposes using a back imitation session to achieve this goal.

Back imitation is the imitation of the learner by the teacher during, before or after the demonstration of a new task. A more complex version of back imitation is mutual imitation in which the demonstrator and the learner take turns imitating each other according to some protocol.

## 2. EXPERIMENTAL DESIGN

The goal of this study was to test the following hypothesis:*H1: Back and mutual imitation conditions will increase the participant's subjective evaluation of the robot's imitative skill as well as her/his intention of future interaction with the robot.*

Thirty six subjects (19 males and 17 females) were recruited from a pool of university students.

During the study, participants interacted with a NAO robot and an agent called WAN. The agent (WAN) is a physically realistic simulation of NAO. After that, subjects filled a pre-experimental questionnaire (PREQ) measuring their background information and expectations for the interaction.

Each participant interacted with NAO and WAN in three sessions (and filled a questionnaire after each session Q1:3). The robot started each session by introducing its name and explaining who will be the leader in this session. The first part of every session was the manipulated part and was conducted in one of in the three following experimental conditions: the No-Imitation (NI) condition in which the leader was WAN (the agent), the Back-Imitation (BI) condition in which the leader was NAO all the time and the Mutual-Imitation (MI) in which the leader was NAO but when the participant failed to copy the pose of the robot, the participant became the leader temporarily until the robot and the participant are in the same pose again (turn-taking).

The second part of each session was started after 20 poses of the leader. This second part of the sessions was identical in all the sessions and only the robot (not the on-screen agent) copied the pose of the subject in real time using the system proposed in [3] with minor modifications.

After finishing the whole experiment the participant filled a post-experiment questionnaire (POSTQ) stating her/his preferences of the three conditions (s)he interacted with.

Five questionnaires were designed for this experiment. The first four (PREQ, Q1, Q2, Q3) measured 22 dependent variables on a Semantic Differential Scale while the POSTQ questionnaire measured the preferences of the subjects on the nine dependent variables corresponding to measurement of robot skill and interaction quality.

The first five questionnaire items (in all questionnaires) measured robot's skill (i.e. accuracy, speed, naturalness of movement, motion human-likeness and overall performance). One item measured participant's self evaluation of her/his imitative skill during the first part of the session. The remaining 16 items were the same used in [4] to measure humanness, shared-reality and likability.

# 3. RESULTS AND DISCUSSION

We divided the 22 evaluation dimensions into six evaluation dimensions. Each one of these dimensions was then analyzed using principal component analysis (PCA) followed by factor rotation using the *normalized varimax* method. This resulted in the number of components for each dimension. Based on this analysis we used intention of future interaction, imitation skill, humanness-positive and humanness-negative, speed, and motion human-likeness as the dependent variables

Each of these dependent variables from questionnaires Q1, Q2, and Q3 were analyzed using factorial t-test and Bonferroni's multiple comparisons correction was applied to counteract the problem of multiple comparisons.

Only *intention of future interaction* showed a statistically significant difference between conditions. MI and BI condition resulted in higher intention of future interaction compared with NI ($p = 0.0293$ and 0.0316 respectively after Bonferroni's correction with effect sizes of 0.435 and .3664 in order) . Likability showed a marginally significant difference between MI and BI conditions (with MI lower than BI) but the difference did not pass the multiple comparisons Bonferroni's test.

To analyze the preferences collected in POSTQ, we calculated a score for every session as follows: If the subject selected one session as best in some dimension, it received a +1 score and if selected as worst it received a -1 score. If only a best session is selected, the remaining two sessions received a -0.5 score and if only a worst session was selected the remaining two session received a +0.5 score. Finally, if no sessions were selected as best or worse (30% of the subjects), the three sessions received a zero score.

A factorial t-test was applied again on all dependent variables of POSTQ using these calculated scores and with the same multiple comparisons and effect size technique described earlier. The only dependent variables that showed statistically significant differences were *Robot's imitative skill* and *motion human likeness* with both MI and BI conditions receiving higher preference scores compared with the NI condition. To further investigate the underlying cause of the difference in subjective evaluation of the robot's imitative skill, we analyzed *all* of the independent variables of POSTQ. There were no statistically significant differences in accuracy, or overall performance. The only independent variable that showed a difference was *naturalness* with both MI and BI conditions receiving higher scores compared with the NI condition. The effect sizes of all of these differences were higher than 0.8 suggesting a strong effect.

These results support the main hypothesis (*H1*). The fact that no differences were directly found in the analysis of Q1:3 may be because the effect is a preference effect between conditions and was not consciously clear to the subjects. Because the goal of the experiment as mentioned in the orientation was to *differentiate* between the three robots, participants' evaluations on the session questionnaires may have been more cognitively mediated that the preferences selected in POSTQ. Given that *intention of future interaction* was higher for the MI and BI conditions compared with NI in session questionnaires may suggest a connection between the subtle preference for these conditions that appeared in POSTQ.

In summary, participants did not show a clear difference in questionnaires Q1:3 regarding the difference in robot's imitative skill between the three experimental conditions (MI, BI, and NI), but a clear difference appeared in their preferences in the post-experimental questionnaire with a strong effect size ($g > 0.8$). The only difference that could be found in session questionnaires (Q1:3) was a slightly higher intention for future interaction for MI and BI conditions compared with NI (with a weak effect size ($g \simeq 0.4$) with a near-significant slight dislike for the MI condition compared with BI.

This superiority of the *BI* and *MI* conditions can be due to the fact that the subject, having to imitate this robot, had to focus more on it. But why would focusing more on the robot cause such an effect? A rigorous and plausible answer of this question requires further investigation but a possible hypothesis is that back and mutual imitation lead to higher anthropomorphism because the participant when imitating the robot has to perceive its motions in a human-like fashion in order to translate them into her/his body form. This may explain why motion human-likeness was enhanced in both experiments by imitating the robot.

# 4. CONCLUSIONS

This paper reported a preliminary study to test the effects of back and mutual imitation on subjective evaluation of imitative skill, interaction quality, humanness of the robot and intention of future interaction. The results show that subjects preferred the imitation of the robot that they previous imitated in terms of imitation skill and reported a higher intention for future interaction with it. This result can be used to inform interaction designers to provide a back imitation session before learning form demonstration sessions.

# 5. REFERENCES


[1] B. D. Argall, S. Chernova, M. Veloso, and B. Browning. A survey of robot learning from demonstration. *Robotics and Autonomous Systems*, 57(5):469–483, 2009.

[2] F. D. Davis, R. P. Bagozzi, and P. R. Warshaw. User acceptance of computer technology: a comparison of two theoretical models. *Management science*, 35(8):982–1003, 1989.

[3] Y. Mohammad and T. Nishida. Tackling the correspondence problem: Closed-form solution for gesture imitation by a humanoid's upper body. In *International Conference on Active Media Technology (AMT)*, pages 84–95. 2013. Maebashi, Japan.

[4] M. Salem, F. Eyssel, K. Rohlfing, S. Kopp, and F. Joublin. To err is human (-like): Effects of robot gesture on perceived anthropomorphism and likability. *International Journal of Social Robotics*, pages 1–11, 2013.