The Answer lies in User Experience: Qualitative Comparison of US and South Korean Perceptions of In-home Robotic Pet Interactions

Ji-Yeong Oh
Department of Intelligence Computing
Hanyang University
Seoul, Korea
ojyhi010309@hanyang.ac.kr

Casey C. Bennett†
Department of Intelligence Computing
Hanyang University
Seoul, Korea
cabennet@hanyang.ac.kr

ABSTRACT

This paper describes a user experience comparison study to explore whether a user's 'cultural background' affects their interaction with in-home pet robots designed for health purposes, e.g. socially-assistive robots (SARs). 11 Koreans and 10 Americans were interviewed after interacting in their own homes with a SAR. Statistical analyses and TF-IDF keyword analyses were conducted to detect significant differences between groups in terms of code co-occurrences. Results showed that American participants were more likely to focus on the interactive experience itself, whereas Korean participants focused more on critiquing technical aspects of the technology. Such differences suggest that Koreans tend to treat robotic pets as "tools", while Americans view the robotic pet through the lens of their past experience raising real-life pets. We discuss implications of this for human-robot interaction (HRI) regarding SARs may be dependent on users' cultural characteristics, e.g. necessitating customized content that takes into account culturally-specific modes of use.

CCS CONCEPTS

Human-centered computing → Human computer interaction (HCI), **Computing Methodologies** → ARTIFICIAL INTELLIGENCE → Natural Language Processing, **Computer systems Organization** → Embedded and cyber-physical systems → Robotics

KEYWORDS

Human Robot Interaction, Socially Assistive Robots, Ecological Momentary Assessment, Robot Sensor Collar, TF-IDF

ACM Reference format:

Ji-Yeong Oh and Casey C. Bennett. 2023. The Answer lies in User Experience: Qualitative Comparison of US and South Korean Perceptions of In-home Robotic Pet Interactions. In *Companion of the 2023 ACM/IEEE Conference on Human-Robot Interaction (HRI'23). ACM, Stockholm, Sweden.* ACM, NY, NY, USA. 4 pages. https://doi.org/10.1145/3568294.3580094

†Corresponding Author

HRI '23 Companion, March 13–16, 2023, Stockholm, Sweden

ACM ISBN 978-1-4503-9970-8/23/03...$ 15.00
https://doi.org/10.1145/3568294.3580094

1 INTRODUCTION

1.1 Background

Robots' expected role in healthcare has been growing noticeably in recent decades. In that capacity, robots can not only serve as an effective tool for in-home health monitoring, but also as a potential tool for treatment, therapy, medication management, compliance, etc. Conversely, patient acceptance of such in-home monitoring robots requires enabling more 'realistic' life-like interactions, allowing them to be viewed as a companion/assistant rather than merely a tool. Past studies have been conducted focused on "user experience-oriented" attributes that may affect the use of the robot. For instance, one such study found that the degree of acceptance of health care robots varies depending on age and gender factors [1]. Another study explored how the personality of a social robot is related to the user's interaction with it [2]. Along those lines, this paper describes an inter-group user experience comparison study based on qualitative data to explore whether a user's 'cultural background' affects their interaction with an in-home health monitoring pet robot. The aim is two-fold: 1) to understand how cultural environment differences influence user perceptions, and 2) to identify other factors should be supplemented in future experiments for improved interaction.

1.2 SAR, EMA App and Robot Sensor Collar

There are three main components that the study participants experienced during the current experiment: socially-assistive robot (SAR) companion pet, an ecological momentary assessment (EMA) mobile app for collecting data about human-robot interactions (HRI) in real-time, and a robot sensor collar for collecting sensor data related to those interactions in user homes. SARs are a common robot platform used for healthcare purposes, such as the seal-like Paro, which has been used as a depression-alleviating companion in older adults [3]. Researchers have shown there exists potential for SARs and its sensor data to ease depression for in-home participating older adults and clinical care staff members. In this study, Joy-For-All robotic therapy pet (https://joyforall.com/, see figure in online Appendix linked below) was used as the SAR deployed into user homes in the United States and South Korea for several weeks at a time, while utilizing the EMA app to collect real-time data about interactions and the user's psychological state [4]. At the end of the deployment, detailed interviews were conducted with participants, the analysis which is the subject of this paper.

2. METHODS

2.1 Interview Coding Methodology

Data analyzed in this paper were post-study interviews conducted by researchers to elicit qualitative feedback about robot companion pet interactions during the 3-week deployment phase in user homes. Interviews were conducted in the participant's native language. The interviews consisted of questions that allowed participants to freely share their experiences with the robot along with the EMA App and sensor collar during the experiment. Questions also prompted users for opinions on future development, such as gamification-related elements (e.g. giving incentives depending on user's interaction amount), sharing data with others, etc. The full list of interview questions can be found in the online Appendix: https://tinyurl.com/4ac58c43

Based on the collected interview scripts, user comments were split into a total of 329 quotes or "documents" from 21 interviews (11 US, 10 Korean). Each quote was then coded using a hierarchical code structure comprising categories developed during previous studies [4,5]. The code structure was initially developed through a thematic analysis of interview data collected during that study to identify common themes participants spoke about. The code group consists of a total of three layers:1) types of technology (SAR, EMA app, sensor collar), 2) detailed features about the technology, 3) individual codes distinguishing positive expressions, negative expressions, and suggestions. The full code structure hierarchy, description of each individual code, and enlarged versions of in-text figures can be found in the online Appendix (link above).

Coding was performed using the Atlas TI software (https://atlasti.com/). To assess the quality of the code hierarchy, inter-rater reliability was calculated by comparing the codings of two independent bilingual coders. We calculated Krippendorff's c-Alpha to be 0.664, which is a tentative result as '$\alpha \geq .667$' is the threshold for moderate agreement.

2.2 Analysis Techniques

Methods are largely divided into three categories: 1) Proximity Operation using Atlas.TI to evaluate co-occurrence frequency, 2) statistical methods for detecting differences in American and Korean interview answers, and 3) natural language processing (NLP) methods via Python NLTK package.

Proximity operators refer to indicators that calculate the spatial relationship between data points (e.g. terms in interview quotes). The 'spatial relationship' can refer to the distance, embeddedness, overlap amount, and the order between data points [6]. These operations allow us to simultaneously represent the relationships between coded data across multiple dimensions. Identifying and analyzing which codes have complex relationships in that way enables interpretation that cannot be examined through analysis of individual codes alone. For example, interviewees typically did not talk only about the specific interview question asked at that moment, but touch on multiple related subjects simultaneously.

Several statistical techniques were also used to detect significant differences in between the two cultural groups. That primarily involved *t*-test (2-tail, independent samples), as well as odds ratios to determine directionality.

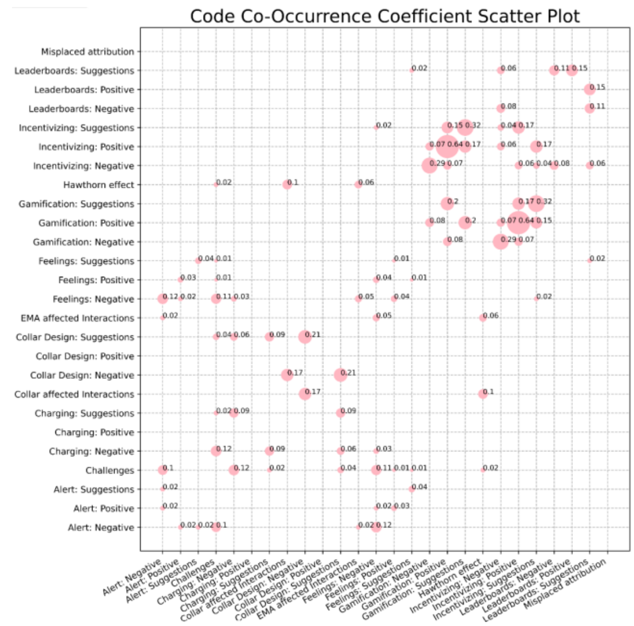

Figure 1. Co-Occurrence C-Coefficient graph. Circle size indicates the strength of the relation between two codes.

Last, we used TF-IDF, an NLP statistical word representation method that weights each word by considering its importance (i.e. keywords) [7]. Weights are calculated by multiplying the Term Frequency (TF) and Reverse Document Frequency (IDF):

$$TF - IDF(t, d, D) = TF(t, d) \times IDF(t, D) \quad (1)$$

TF is the frequency of appearance of a particular word *t* in a particular document *d*. IDF is the reciprocal of the Document Frequency (DF) value, where DF is the number of documents with a specific word *t* appearing in the entire document *D*. Based on that, cosine similarity between documents can be calculated

3. RESULTS AND DISCUSSIONS

3.1 Co-Occurrence Analysis

The co-occurence graph indicating the strength of the relation between each pair of codes is shown in **Figure 1**. Below we summarize insights from that regarding improvements for better interaction between in-home robotic pets and humans.

3.1.1. Rewards as Interaction Themselves. Among the interview questions, there were question prompts asking the participants' views on the addition of 'gamification elements' (e.g. turning the interaction into a "competition", leaderboards, etc.) to the EMA App to foster more interaction with the robotic pet. Similarly, there were additional questions regarding mechanisms in which participants get incentives based on their interactive behaviors with the SAR companion pet. Interestingly, there seemed to be a close connection between participants views on gamification and incentives. The co-efficient between 'Gamification: Positive' and 'Incentivizing: Positive' was 0.64, followed by 0.32 (between 'Gamification: Negative' and 'Incentivizing: Negative') and 0.29 (between 'Gamification: Suggestions' and 'Incentivizing: Suggestions'). This pattern can be seen in some quotes, for instance one Korean participant said:

code name	US Average	KR Average	US Std	KR Std	T-Test value	p-value	odds ratio
Collar	1.7000	2.6364	0.9487	1.4334	-1.7798	0.0925	0.5107
EMA App	7.6000	7.5455	2.1187	2.2074	0.0578	0.9545	0.7748
Overall Experiment	8.7000	5.3636	3.0569	1.6293	3.0769	0.0085	1.7758

Table 1. T-Test & Odds Ratio Table (code group 1)

code name	US Average	KR Average	US Std	KR Std	T-Test value	p-value	odds ratio
Alert	3.000	3.818	1.247	1.601	-1.3125	0.205	0.629
Challenges	1.600	2.091	0.699	1.375	-1.0447	0.313	0.640
Charging	1.100	1.636	0.994	0.674	-1.4324	0.172	0.563
Design	0.700	0.636	0.823	0.924	0.1669	0.869	0.962
Feelings	7.100	3.455	3.315	1.508	3.1905	0.008	2.277
Gamifications	0.600	1.000	0.843	0.775	-1.1284	0.274	0.510
Incentivizing	1.700	1.455	1.160	0.522	0.6151	0.550	1.026
Interactions	2.200	1.727	1.135	1.489	0.8222	0.421	1.130
Leaderboard	1.000	0.818	0.816	0.603	0.5758	0.573	1.074

Table 2. T-Test & Odds Ratio Table (code group 2)

- *"It's a kind of interaction, like growing a character, I hope that certain functions would be revealed and the incentives would make the interaction level up to the higher state. If it can be realized, I would communicate with more fun."*

Some participants also suggested giving rewards for 'interaction with the robot itself', e,g, raising game characters and opening up an 'Easter Egg' upon certain milestones. These suggestions can be interpreted as reflecting the desire for more meaningful long-term interaction with robots that participants potentially can have. This desire is very similar to the expectations of human relationship [8,9]. The attitude of expecting a robot to give them a richer and more diverse response as a reward when they interact well with it, is closely associated with the attitude of expecting the other human to concentrate on their behavior and react continuously in meaningful ways. More broadly, that suggests it may be useful to diversify a robot's range of behavior and allow the robot to modify its own long-term behavioral responses.

3.1.2. User Friendly Design. The codes [Collar Design: Suggestions] and [Collar Design: Negative] also co-occurred frequently, which seemed to relate to a significant number of comments on the robot collar's rugged design, and suggested changing the design to a more user-friendly and sophisticated shape (c-coefficient 0.21). Interestingly though, there was no question directly asking the participants how he or she liked the collar's design during the interview. For example:

- *"And since it is not completely final version, it was kind of rough. If you make it prettier… (it would be great)"*
- *"I think it'd be good to make the Collar of the sensor into a form of cat's leash. Yeah, it was like a machine."*
- *"The collar was bulky. That was about it. Moving it was more annoying, if it was smaller that would be more helpful, but I know there's constraints."*

The feedback indicated that the current collar design is too machine-like and bulky to use in a natural manner, which is significant as the robot's physical design directly or indirectly influences the user's experience with the robot. These findings suggest focusing more on user comfort rather purely on functionality is critical for SAR companion pets.

3.1.3. Emotional Responses to Robot Pets. Four codes (or code groups) met the maximum significant p-value criterion of '≤ 0.05': [Overall Experiment], [Feelings], [Feelings: Negative] and [Feelings: Positive]. Based on *t*-test comparison, it was clear that American participants had more comments about 'feelings' (both negative and positive emotions) than Koreans. Looking at the quotations of U.S. participants on 'Positive' emotions, many of them said specifically that they, their acquaintances, and their families felt comfort and joy while interacting with the robot.

- *"I enjoyed it, I like, I'd come home from my class or something I look forward to be able to interact with the dog. Just being able to play with it and talk to it"*
- *"I think that this is great, um, in my experience, because I took it home and I was at my grandma's house and she actually liked it and she claims that she doesn't like dogs or pets. But she enjoyed interacting with it, and I think it's because her and her husband retired, so they just sit at home, all day ... that was like kind of a highlight that I enjoyed."*
- *"Honestly, the study just kind of brought back memories and like when I got my puppy, so it just it felt good to like reflect on that. or just to like look forward to having like something to come home to I guess. So yeah."*

The US quotations related to 'Negative' emotions, interestingly, were often about their 'guilt' or 'irritation' when EMA notifications appeared but they could not interact with the pet robot for inevitable reasons.

- *"Um, mm. Yeah, sometimes I did feel a little guilty at times when like I couldn't just bring it out cause of it. But most of the time I was just able to, when I saw it, it was just that reminder like, oh, let me pull it out."*
- *"Yeah a little bit because I will feel a little guilty if it is, would you please interact with it so it's like okay yeah I'm gonna interact with it."*
- *"I mean it was annoying when I wasn't at home, and you know couldn't interact with it, you know it's just like any other notification."*

On the other hand, Korean focused more on the 'technology' itself rather than the 'interaction' with the robot, and their interview data generally lacked comments about a sense of guilt. One possible reason why participants from the two different cultures showed different reactions could be related to past experiences related to "intimacy with pets". In general, Koreans have a significantly lower degree of familiarity with pet animals

Code Group	Korean - Korean	American - American	Korean - American
Entire Interview	0.255	0.369	0.137
Overall Experiment	0.246	0.401	0.160
Feelings	0.307	0.455	0.203
Feelings: Positive	0.270	0.446	0.184
Feelings: Negative	0.304	0.377	0.157

Table 3. Average TF-IDF Cosine Similarity Within and Between Culture, by Code Group

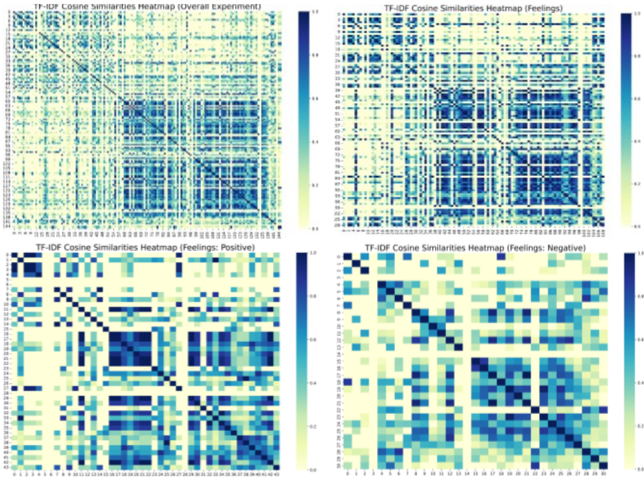

Figure 2. TF-IDF Cosine Similarity Heatmap among Documents by Unit of Code (see Table 3)

than Americans, which has been found to affect robotic pet interactions in previous studies [10]. This may be due to Korea's unique residential environment, which has a very high population density compared to the United States (over 42,500 people per square mile in the Seoul area). It is common to live in tightly-packed, single-person residential spaces (called a 'one-room village') in Korea, where there is not enough space to raise pets and where soundproofing in homes is minimal. As such, the proportion of Korean households raising pets is lower than that of the United States – 15% vs. 67% of all households nationwide as of 2020 [11,12]. Putting these together, one can infer that differences detected in this study underscore how Korean participants treat robot pets more as an experimental tool, while Americans perceived the robot as an interaction partner or even a substitute for pets who they previously raised. This suggests that differences in cultural environment can significantly affect their interaction with robotic pets.

3.2 NLP Analysis via TF-IDF

Cosine similarity between individual documents (i.e. quotations derived from TF-IDF vectors corresponding to code groups) showed significantly different code distributions between the two cultures, as well as interview overall. Their average similarities grouped by document groups (i.e. code groups) are described in Table 3. For brevity, significant results can be seen in Figure 2 in a form of heatmaps. Significant areas of correlation are represented by the darker rectangular patches of blue, i.e. the heatmaps show significant clustering for certain code groups. Table 3 shows some of those code groups broken out by culture.

Figure 3. TF-IDF Top50 Keywords Word Cloud (left: Korean, right: American)

As the table shows, average values of TF-IDF cosine similarity between different cultures (Korean-American) were lower than those between the same culture (Korean-Korean and American-American). In other words, cosine similarity suggested the two groups clustered together based on the interview quote content.

Given the above, we then compared the top keywords of each document group. Results are shown in Figure 3. Keywords between the two groups exhibit notable differences, which can be summarized as Koreans focused on the technology itself (e.g. EMA notifications), while Americans focused on the interactive experience with the robot. That further supports our conclusion that Koreans treated the robotic pet as an experimental tool, while Americans viewed it more as an interactive evocative object.

4. DISCUSSION

In summary, the key takeaway for HRI from this study is the differing perceptions of in-home robotic pets based on a user's cultural environment. Even if the appearance and behavior of the robot is the same, the interaction result may be different from designed expectations based on users' home living environments and their past experience raising pets. This suggests that a future strategy in HRI may need to better consider cultural factors that impact particular robot applications, e.g. robot pets.

Understanding *user experience* as a key part of interaction with robots should not be overlooked when designing HRI systems, as long as successful interaction demands understanding particular modes of use [4,13,14]. Based on post-interviews of participants here using in-home robot pet, this paper analyzed how users perceived their experiences associated with such an HRI system: an interaction monitoring tool (EMA App), robot sensor collar, and communication with the robot itself. Statistical and NLP techniques were used to find out whether there was a significant difference between Korean and American user experiences as well as quantify the specific ways that they were different. Results showed that Koreans viewed pet robots as experimental tools due to cultural environment factors and lack of familiarity raising pets, while Americans viewed the robot as the subject of interaction and as a companion.

ACKOWLEDGEMENTS

This work was supported by funding provided via the Hanyang University Research Fund, HY-2020 (in Korea), and from the National Science Foundation grant, IIS-1900683 (in USA). We would also like to thank our research collaborators.

REFERENCES

[1] I.H. Kuo, J.M. Rabindran, E. Broadbent, Y.I. Lee, N. Kerse, R. M. Stafford and B. A. MacDonald, 2009. Age and gender factors in user acceptance of healthcare robots. In *RO-MAN 2009-The 18th IEEE International Symposium on Robot and Human Interactive Communication.* IEEE, 214-219. DOI: https://doi.org/10.1109/roman.2009.5326292

[2] B. Tay, Y. Jung and T. Park, 2014. When stereotypes meet robots: the double-edge sword of robot gender and personality in human–robot interaction. *Computers in Human Behavior*, *38*, 75-84. DOI: https://doi.org/10.1016/j.chb.2014.05.014

[3] N. Randall, C.C. Bennett, S. Šabanović, S. Nagata, L. Eldridge, S. Collins and J.A. Piatt, 2019. More than just friends: in-home use and design recommendations for sensing socially assistive robots (SARs) by older adults with depression. *Paladyn, Journal of Behavioral Robotics*, *10*(1), 237-255. DOI: https://doi.org/10.1515/pjbr-2019-0020

[4] C.C. Bennett, Č. Stanojević, S. Šabanović, J.A. Piatt and S. Kim, 2021. When no one is watching: Ecological momentary assessment to understand situated social robot use in healthcare. In *Proceedings of the 9th International Conference on Human-Agent Interaction.* 245-251. DOI: https://doi.org/10.1145/3472307.3484670

[5] C.C. Bennett, S. Sabanovic, J.A. Piatt, S. Nagata, L. Eldridge, and N. Randall. 2017. A robot a day keeps the blues away. IEEE International Conference on Healthcare Informatics (ICHI), 536-540. DOI: https://doi.org/10.1109/ichi.2017.43.

[6] T.M. Paulus and J. N. Lester, 2016. ATLAS TI for conversation and discourse analysis studies. *International Journal of Social Research Methodology*, 19(4), 405-428. DOI: https://doi.org/10.1080/13645579.2015.1021949

[7] A. Aizawa, 2003. An information-theoretic perspective of TF–IDF measures. Information Processing & Management, 39(1), 45-65. DOI: https://doi.org/10.1016/s0306-4573(02)00021-3

[8] D. Martinho, J. Carneiro, J.M. Corchado and G. Marreiros, 2020. A systematic review of gamification techniques applied to elderly care. *Artificial Intelligence Review*, *53*(7), 4863-4901. DOI: https://doi.org/10.1007/s10462-020-09809-6

[9] N.L. Robinson, S. Turkay, L.A. Cooper and D. Johnson, 2019. Social robots with gamification principles to increase long-term user interaction. In *Proceedings of the 31st Australian Conference on Human-Computer-Interaction*, 359-363. DOI: https://doi.org/10.1145/3369457.3369494

[10] C.C. Bennett, C. Stanojevic, S. Kim, S. Sabanovic, J. Lee, and J.A. Piatt, 2022. Comparison of In-home Robotic Companion Pet Use in South Korea and the United States: A Case Study. 2022 9th IEEE RAS/EMBS International Conference for Biomedical Robotics and Biomechatronics (BioRob). IEEE, 01-07. DOI: https://doi.org/10.1109/biorob52689.2022.9925468

[11] Korean Statistical Information Service, KOSIS, 2020. https://url.kr/3v4m6u

[12] American Pet Products Association, APPA, 2021. https://globalpetindustry.com/article/appa-survey-reveals-trends-us-pet-ownership-spending-shopping-habits-pet-acquisition-and

[13] P. Flandorfer, 2012. Population ageing and socially assistive robots for elderly persons: the importance of sociodemographic factors for user acceptance. *International Journal of Population Research*, *2012.* DOI: https://doi.org/10.1155/2012/829835

[14] R.A. Søraa, G. Tøndel, M.W. Kharas and J.A. Serrano, 2022. What do Older Adults Want from Social Robots? A Qualitative Research Approach to Human-Robot Interaction (HRI) Studies. *International Journal of Social Robotics*, 1-14. DOI: https://doi.org/10.1007/s12369-022-00914-wPrice:$15.00