ISASE 2019

Intrinsic Motivation in Virtual Assistant Interaction

Chang LI* and Hideyoshi YANAGISAWA **

** The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0013, Japan*
lichang@mail.design.t.u-tokyo.ac.jp
*** The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0013, Japan*
hide@mail.design.t.u-tokyo.ac.jp

Abstract: Conversational virtual assistants have exploded in popularity in recent years. Developers have been working to create personality in virtual assistants, so that users treat them more than mere automation tools. Among other factors, motivation is considered playing an important role. However, previous researches focused mainly on extrinsic motivation, and ignored the fact that intrinsically motivated engagement cultivates better relationship of reliance. Intrinsic motivation is derived from inner urge to seek enjoyment in an action itself, while extrinsic motivation is oriented at consequences and rewards. In context of virtual assistant interaction, motivation is formed based on the user's prior expectation: whether the assistant can cope with his/her intention. When the expectation is contradicted, the user is subjected with disruption and is bounced back to learning process about the assistant's attributes. Related works in expectation effect found that the surprising amplitude of disruption is function of perception uncertainty and prediction error. The current research investigated effects of expectation towards the assistant's capability and uncertainty during perception on intrinsic motivation. We conducted laboratory environment experiments using Amazon Echo smart speakers and extracted subjective assessment as well as used free choice paradigm as objective measurement. The result showed a significant trend that small uncertainty encouraged more intrinsically motivated interaction while large uncertainty motivated the participant to interact in order to resolve uncertainty. Neither uncertainty nor expectation showed main effect on subjective assessment on intrinsic motivation. Our findings suggest that trading consistency for performance is a risky strategy in designing virtual personality.
Keywords: *Intrinsic Motivation, Virtual Assistant Interaction, Expectation Effects, Smart Speaker, Wizard of Oz Experiment*

1. INTRODUCTION

1.1 Background

Virtual assistants prevail in today's smart product market. Likes of Apple Siri, Amazon Alexa and Google Assistant that reside in a smart speaker allow users to merely speak to get commonplace household chores done, such as controlling room lighting, checking weather and news. Moreover, they were also developed to amuse people: that is about game and pastime functions. From basic games such as rock paper scissors and Simon Says, to trivia and quiz.

Beyond creating a mere service of automation, developers in the field have moved on to create a personality, one that is proactive and fun to interact with. It has been reported that kids are willing to befriend with conversational assistants in household; however, adults are in general conservative towards friendship with a virtual assistant [1]. Motivation towards interaction has recognized as important factors in building a virtual personality [2].

However, studies on human-robot interaction treated motivation as a simple factor and failed to distinguish intrinsic motivation from extrinsic motivation. One of the previous studies typically used self-reported willingness to carry out tasks as indicator of participants' motivation, despite the fact that carrying out any task should be regarded precisely extrinsic motivation [3]. Intrinsic motivation towards interaction remained poorly studied. There has been proof that intrinsic motivation plays a much more important role on enhancing engagement quality than extrinsic motivation [4]. The current research aims at investigating factors of intrinsic motivation in context of virtual assistant interaction.

By nature, intrinsic motivation is derived from the expectation of enjoyment through taking an action. On the other hand, extrinsic motivation is oriented at consequences of the action. Among established theories on hedonic experiences, flow experience is highly transferable for virtual assistant interaction [5]. Moderate level of challenge is necessary to encourage interaction, while the user must not be overwhelmed by the challenge or feel unconfident in the assistant and himself to cope with it. The outcome of single conversation can be categorized either expected or unexpected. As for human-computer interaction, unexpectedness is often negative; when disruption occurs, the user has to relearn about the artifact, which harms motivation and hampers

engagement [6].

According to theories of expectation effects, unexpectedness (termed "surprise" in related literatures) is function of uncertainty, prediction error and noise [7]. It has been found that uncertainty diminishes expectation effect, i.e. the perceptual bias. In context of virtual assistant interaction, prediction error can be interpreted as the gap between expectation of capability (e.g. smartness) and perceived capability after a conversation resolves. A question arises: is user's intrinsic motivation more sensitive to a) an assistant of high capability expectation paired with large uncertainty; or b) an assistant of low capability expectation paired with small uncertainty?

1.2 Objectives

The main objective of the current research is to discover effects of uncertainty and capability expectation on intrinsic motivation. In context of virtual assistant interaction, uncertainty can potentially affect interaction strategy. Given a relatively low capability expectation, moderate level of uncertainty leaves the user with expectancy for potential greater-than-usual performance. On the other hand, uncertainty also means risk to fail to cope with seemingly easy commands. Moreover, capability expectation must also be considered a dependent factor of any type of motivation towards interaction. High capability expectation implies positive assessment of smartness, thus affecting motivation.

3. METHOD

For the above stated objectives, we conducted experiment using the smart speaker Amazon Echo, whose virtual assistant is named Alexa. Participants were asked to interact with the virtual assistant and carry out tasks that are designated by task lists. Based on previous studies and established theories, we propose the following hypotheses:

1. *Effects on intrinsic motivation.* Uncertainty and capability expectation will have interaction effects over intrinsic motivation. Intrinsic motivation will increase with expectation increment when uncertainty is small; and decrease when uncertainty is large.
2. *Effects on interface transparency.* Uncertainty and capability expectation will also have interaction effects on interface transparency, i.e., the extent to which the assistant's computing process is perceived as understandable. Interface transparency will increase with expectation increment when uncertainty is small; and decrease when uncertainty is large.

To verify the hypotheses, we manipulated the virtual assistant's responses using the Wizard of Oz's Method, in order to shape learning processes thus forming expectation towards capability and uncertainty about the capability. The advantage of using a smart speaker is that it meets the requirement of Free Choice Paradigm, an objective measurement of intrinsic motivation, due to its vast range of interactive options from trivia to games.

3.1 Applying Wizard of Oz method

We used Wizard of Oz Method to manipulate the virtual assistant's response. This was achieved by streaming response sound tracks. The responses were generated using the Echo simulator provided by Amazon Developer.

The experiment was conducted in a room where the participant sits half a meter from a table, upon which the smart speakers were placed. A dense black curtain hid the smart speakers from the participant in order to hide the LED ring on top of the smart speakers, which would otherwise spoil experimental fidelity because the smart speakers were actually muted during task sections, glowing red. The experimenter sits at a desk approximately two meters behind the participant, able to hear the conversation clear while not distracting the participant.

3.2 Manipulation of expectation

Participants were divided into two groups by uncertainty level. Expectation levels were arranged within-group, so that each participant took one low expectation section and one high expectation section respectively. The expectation is manipulated through task complexity and the assistants' ability to comply, as elaborated in Table 1.

Table 1: Manipulation by section

Exemplary tasks	Low expectation section	High expectation section
Simple tasks	able	able
Superposition of simple tasks	able	able
Connective tasks (A and then B)	unable	able
Memorizing variables	unable	able
Calculation using variables	unable	unable

Participants were made to believe that they were interacting with a distinct character in each section. It is presumed that they think of one character smarter and generally more capable than the other one, as result of manipulation. To counter order effect, the experiments are planned so that half of the participants first undergoes low expectation section, and the section order is reversed for the other half of participants.

3.3 Manipulation of uncertainty

It is presumed that, if the contradiction between the assistant's performance and the participant's expectation remains unresolved at the end of the section, he tends to

feel uncertain about the assistant's capability.

To raise uncertainty, the assistants' behavior in high uncertainty group was manipulated in such a way that it would: 1) fail at a simpler task given that it had successfully carried out another task of similar or higher complexity; and 2) accomplish a complex task, given that it had failed at similar or simpler tasks. Though designated with the same lists of tasks, the virtual assistant in small uncertainty group showed a clear profile of capability, whereas in large uncertainty group its behavior would contradict with the participant's expectation.

3.4 Measurement

As for subjective measurement we used a questionnaire which extracts assessment (7-point Likert scale) on five sub scales: Interest/Enjoyment, Relatedness, Smartness/Capability, Anthropomorphous Features and Interface Transparency. The first two were derived from the Intrinsic Motivation Inventory [4], which are solely responsible for intrinsic motivation and the extent to which our virtual assistant is trusted and liked. The other three sub scales were included to cover positive attributes for smart product design, especially those involving interaction with an avatar. After each task section, the participant is asked to answer the questionnaire. In addition, free choice paradigm was also used as objective measurement. After the second set of questionnaires are filled out, the experimenter would pretend that he has to print a final survey document, so that the participant is left alone in the room. In order to record the participant's behavior during the free choice period, the native logging function of Amazon Echo is supplemented by a hidden smartphone left well-hidden on the experimenter's desk.

3.5 Participation

Six participants (1 female, 5 male) took part in the experiment. Participants were students at the University of Tokyo and were recruited using personal connection. All of the participants met the criterion that they had little experience with conversational virtual assistant, nor own smart speaker of any type. Participants were randomly assigned to one the uncertainty group.

4. RESULTS

Since assessment data from both groups showed normal distribution (Shapiro–Wilk test), two-way ANOVA was used to test effects of uncertainty and expectation.

4.1 Effect on intrinsic motivation

The questionnaire scores did not show significant effect by uncertainty (F=0.917; p=0.392) nor by expectation (F=0.5; p=0.519). Figure 1, the intrinsic motivation versus expectation plot featured a pair of parallel lines. Test for interaction effect contradicted with our hypothesis 1 (F=0; p=1.000).

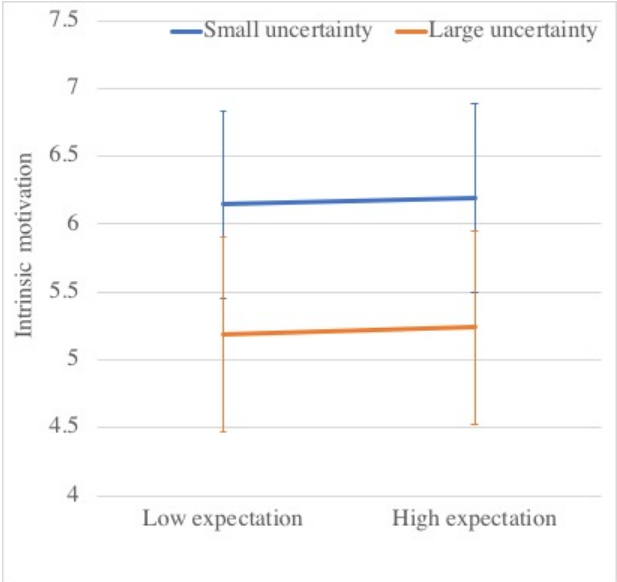

Figure 1: Effect on intrinsic motivation

However, subjective measurement showed a significant trend that large uncertainty increased intrinsic motivation (see Table.2). Number of initiations by the participant during the free choice period was chosen as dependent variable.

Table 2: Number of initiations

Participant No.	Uncertainty group	Pastime activities	Epistemic activities
1	Small	0	0
2	Small	0	0
3	Small	3	0
4	Large	0	5
5	Large	0	0
6	Large	6	3

The reason not to use interaction duration is that it highly depends on the nature of interaction. All participants targeted the virtual assistant of high expectation section, regardless with section order and regardless with uncertainty group. Paired T-test resulted in p=0.0934, a significant trend that participants in large uncertainty group had more initiations compared with small uncertainty group.

4.2 Effect on interface transparency

Two-way ANOVA on the Interface Transparency sub scale showed a significant trend for interaction effect of uncertainty and expectation (F=3.464; p=0.136).

As shown in Figure 2, in small uncertainty group, interface transparency was reported higher for the high expectation section compared with that for the low expectation section. In large uncertainty group, interface transparency was reported lower for the high expectation section.

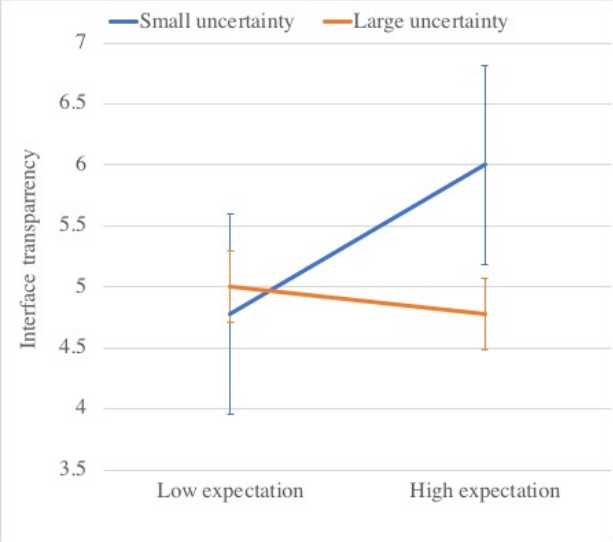

Figure 2: Effect on interface transparency

5. DISCUSSION

Both hypotheses were not supported by the experiment result with the six participants since no significant main effect nor significant interaction effect were found. Subjective measurement for intrinsic motivation contradicted with hypothesis 1 with a non-significant trend ($p=0.0934$). For hypothesis 2, p-value for interaction effect test scored 0.136. We expect a significant interaction effect after more participants are involved.

5.1 Free choice initiation by motivation types

Results of objective measurement indicated a significant trend which contradicts hypothesis 1. However, number of initiations varied drastically by individual, as can be seen in Table 2, making less convincing of this finding.

Interaction that happened during free choice period can be divided into two categories: pastime activity and epistemic activity. These derive from intrinsic motivation and experiential motivation, respectively. Pastime included song requests, casual chat, etc. while epistemic activity typically took form of retrying a failed task during the experiment sections (refer Table 3 for detail).

We noticed a trend despite sparse data, that type of motivation behind initiation differed across groups. In large uncertainty group, 7 out of total 14 interaction records were aimed at testing capability of the assistant. Two participants retried three and two tasks from the task list, respectively. On the other hand, all initiation records in small uncertainty group were pastime and playing request. Studies on motivation taxonomy categorized playing and pastime activities as intrinsically motivated, whereas epistemic curiosity is regarded experiential motivation, a stance between extrinsic and intrinsic motivation [8]. Participants who experienced great uncertainty would reasonably desire to explore the assistant's capability in order to resolve uncertainty. If we could reach a large enough sample size, we expect a significant split in motivation type between uncertainty group.

Table 3: Interaction record in free choice period

Participant No.	Uncertainty group	Speech
3	Small	Alexa, are you awake?
		Alexa, it was very interesting to talk to you.
		Alexa, can you say hello in Japanese?
5	Large	Alexa, roll a die and flip that many coins.
		Alexa, roll a die.
		Alexa, flip three coins.
		Alexa, my name is Tim.
		Alexa, who am I?
6	Large	Alexa, sing a song.
		Alexa, play another one.
		Alexa, play a song.
		Alexa, sing a song.
		Alexa, turn the light yellow.
		Alexa, turn the light off.
		Alexa, sing another song.
		Alexa, change another one.
		Alexa, one plus one?

5.2 Bias induced by forgiveness

In the light of hypothesis 1, we expected that in large uncertainty group, intrinsic motivation should be apparently lower when capability expectation is high. Our results showed no reaction of self-reported intrinsic motivation to expectation.

This could be interpreted as a buffering effect caused by participants' forgiveness towards virtual assistant with artificial intelligence. Previous studies on Human-Robot Interaction [3] found that erratic behavior of a robot had limited negative impact. It lowered the participants' subjective assessment, but it did not impact self-reported willingness to comply with the robot's unusual requests. The authors mentioned that participants were being "forgiving" to robots.

This assumption aligned with our findings. When asked for feedback, our participants (3 out of 6 participants, 2 from large uncertainty group) remarked highly the virtual assistants' humanlike talking skills, specifically in the task scenario where the participant would tell it to remember information, and it would asked them to confirm by saying "You want me to remember …, right?". When participants confirmed with "Yes", the virtual assistant said "Ok. Got it.", which gave them a feeling of natural, back and forth conversation. These could account for their forgiveness on the assistant's failures elsewhere during the experiment.

On the other hand, our participants reported being

forgiving, typically saying virtual assistants are “after all not very all-around” (2 of 6, both from large uncertainty group), and “such failure are understandable; the tasks would be difficult for other assistants as well” (1 out of 6). The two participants whose feedback implied forgiveness exclusively scored higher intrinsic motivation score than the only other participant left in large uncertainty group.

6. CONCLUSION

Our study found significant trend of interaction effect by uncertainty and capability expectation of a virtual assistant on its interface transparency. When uncertain about capability, self-reported interface transparency decreases when expectation increases.

In addition, we found that uncertainty encouraged the participants to interact in the hope to examine the assistant’s capability, in other words, to explore system boundaries. On the other hand, those who experienced small uncertainty tend to be intrinsically motivated into interaction.

Our current findings suggest that prioritizing singular high performance over consistency is a risky design strategy. Raising performance did not improve intrinsic motivation unless uncertainty was decreased (see Figure 1). In ideal case, computation models based on confidence threshold should be customizable to users’ acceptance of inconsistency, because users with long term using experience might appreciate high performance despite inconsistent performance.

Our research complements previous human-robot interaction researches, in that we covered intrinsically motivated activities as subjective measurement. A virtual assistant serves more purpose than a mere automaton, therefore insights on intrinsic motivation is of the essence.

The major limitation of current research is that it involved only laboratory-environment-experiments, which allows mere an hour of interaction time at best. Individual differences regarding technology acceptance could be mitigated if the participants are given adequate time to settle their expectation towards the virtual assistants. Forgiveness bias can also be eliminated when the novelty gradually wears off. As for future work, long term engagement should be studied with focus on emotion and motivation type behind interaction.

7. REFFERENCES

[1] Sciuto, A., Saini, A., Forlizzi, J., & Hong, J. I. (2018, June). Hey Alexa, What's Up?: A Mixed-Methods Studies of In-Home Conversational Agent Usage. In Proceedings of the 2018 on Designing Interactive Systems Conference 2018 (pp. 857-868). ACM.

[2] Robinson, N., Connolly, J., Johnson, G., Kim, V., Hides, L. & Kavanagh, D. (2018). Measures of incentives and confidence in using a social robot. Science Robotics. 3. eaat6963. 10.1126/scirobotics.aat6963.

[3] Salem, M., Lakatos, G., Amirabdollahian, F., & Dautenhahn, K. (2015, March). Would you trust a (faulty) robot?: Effects of error, task type and personality on human-robot cooperation and trust. In Proceedings of the Tenth Annual ACM/IEEE International Conference on Human-Robot Interaction (pp. 141-148). ACM.

[4] Ryan, R. M., & Deci, E. L. (2000). Intrinsic and extrinsic motivations: Classic definitions and new directions. Contemporary educational psychology, 25(1), 54-67.

[5] Csikszentmihalyi, M. (2014). Toward a psychology of optimal experience. In *Flow and the foundations of positive psychology*(pp. 209-226). Springer, Dordrecht.

[6] Krippendorff, K. (2005). The semantic turn: A new foundation for design. crc Press.

[7] Yanagisawa H, Kawamata O and Ueda K (2019) Modeling Emotions Associated With Novelty at Variable Uncertainty Levels: A Bayesian Approach. *Front. Comput. Neurosci*. 13:2. doi: 10.3389/fncom.2019.00002

[8] Banyte, J., & Gadeikiene, A. (2015). The effect of consumer motivation to play games on video game-playing engagement. Procedia Economics and Finance, 26, 505-514.