Human Factors in Design, Engineering, and Computing, Vol. 238, 2026, 2864–2872
https://doi.org/10.54941/ahfe1008510

Effects of Avatar-Mediated Helping and Gratitude Feedback on Users' Prosocial Behavioral Intentions

Akihiro Ogino

Faculty of Information Science and Engineering, Kyoto Sangyo University, Kyoto, Japan

ABSTRACT

Helping others and receiving gratitude may promote positive affect, well-being, and prosocial behavior. However, interpersonal helping may be inhibited by fear of rejection or uncertainty about receiving appreciation from others. We developed a web application that framed solving Sudoku puzzles as helping an in-app avatar and examined whether visual avatar responses and verbal gratitude affected prosocial behavioral intentions, affective states, depressive symptoms, avatar attachment, and application use. Thirty university students (26 men and 4 women; mean age = 21.63 years) completed a within-participant study with three counterbalanced 12-day conditions: Task-Only, Avatar-Response, and Avatar-Gratitude. The mean changes in prosocial behavioral intention from baseline to Day 12 were –1.467, 1.700, and 3.633, respectively. Although the omnibus Friedman test did not reach significance ($p = .089$), the prespecified within-participant comparison showed a greater increase in the Avatar-Gratitude condition than in the Task-Only condition after the Holm correction (adjusted $p = .042$, Cohen's $d = 0.53$). The Avatar-Gratitude change was also greater than zero after correction (adjusted $p = .015$). No significant difference emerged between the two avatar conditions, and no consistent condition effects were found for affect, depressive symptoms, attachment, use, or self-reported effectiveness. Thus, combining avatar responses with gratitude may strengthen prosocial behavioral intentions relative to presenting a task without an avatar, but the independent contribution of verbal gratitude has not been established. Because participants used each application only about four times on average, avatar-based interventions may require immediate, salient, and individually adaptive feedback from the beginning.

Keywords: Affective computing/kansei engineering, Social computing, Human-avatar interaction, Gratitude feedback, Prosocial behavioral intentions

INTRODUCTION

Prosocial behaviors can create positive experiences for recipients and helpers (Hu et al., 2016). Positive computing identifies gratitude and contribution to others as relevant design elements for technologies intended to support human well-being (Calvo and Peters, 2014). Receiving gratitude after helping may communicate that one's contribution has social value and encourages subsequent prosocial behavior (Bartlett and DeSteno, 2006; Grant and Gino, 2010).

Received September 15, 2026; Revised October 15, 2026; Accepted October 26, 2026; Available online December 1, 2026

© 2026 The Authors. This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 License.
For more information, see https://creativecommons.org/licenses/by-nc-nd/4.0/

Interpersonal helping nevertheless involves barriers. People may fear that an offer will be rejected, worry that it will burden the recipient, or remain uncertain whether their help will receive any response. A virtual avatar that consistently responds to a user's action may reduce some of these barriers and provide a controlled setting in which helping can be experienced. Research on virtual agents and gamified applications indicates that avatar appearance and feedback design can affect motivation and intentions to reuse a system (Baylor, 2009; Jahn et al., 2021). However, few longitudinal studies have separated an avatar's presence and visual response from feedback that also includes verbal gratitude when examining prosocial behavioral intentions.

This study examined whether framing task completion as helping an avatar, together with avatar responses and gratitude feedback, influenced users' prosocial behavioral intentions, affective states, depressive symptoms, attachment, and application use. The prespecified primary hypothesis was that Avatar-Gratitude would produce a greater increase in prosocial behavioral intention than Task-Only. The Avatar-Response versus Avatar-Gratitude comparison examined the incremental contribution of verbal gratitude to the overall gratitude felt by participants.

PROPOSED SYSTEM

Task Design as Helping Behavior

We developed a web application to create a situation in which users spent their time helping an avatar by solving Sudoku puzzles (Figure 1). Participants could access the application from their smartphones at any time. After completing a puzzle, they were automatically returned to the avatar screen, where a visual response and, depending on the condition, a gratitude message appeared. For the two avatar conditions, the instructions explicitly described the avatar as being in need of help and the Sudoku task as a means of helping the avatar.

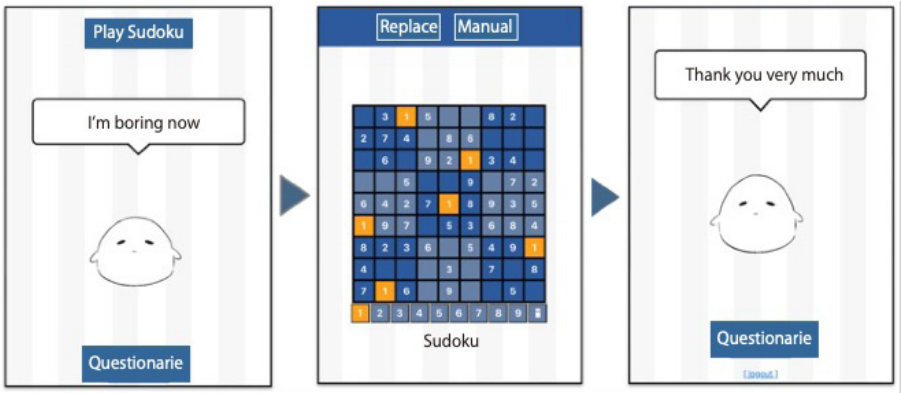


Figure 1: Application flow.

Affinity and Gratitude Feedback

The avatar was displayed as a GIF animation. Its movement, facial expression, and visual effects changed across 15 affinity levels calculated from cumulative application use. Five levels of facial expression and movement were combined with effects such as musical notes, blushing cheeks, and stars to visualize accumulated helping. In the gratitude condition, messages of appreciation also changed as affinity increased, and one of several messages was randomly selected within each level. This variation was intended to reduce the mechanical impression created by repeated wording and to convey that the user's relationship with the avatar was changing. Accordingly, Avatar-Gratitude versus Task-Only evaluates the combined feedback package, whereas Avatar-Gratitude versus Avatar-Response tests the incremental contribution of verbal gratitude.

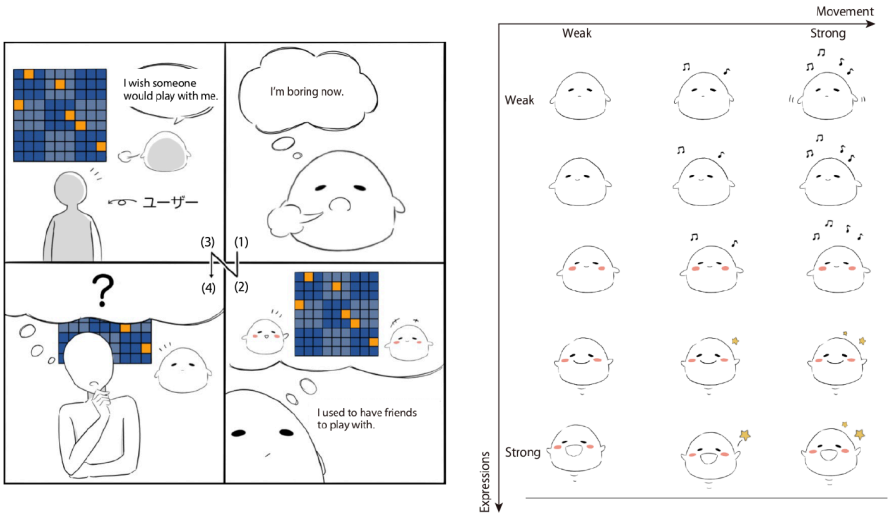


Figure 2: Introductory comic used in the avatar conditions and avatar presentation across 15 affinity levels.

EXPERIMENT

Participants and Experimental Design

Thirty university students (26 men and 4 women; mean age = 21.63 years) participated. They received an explanation of the study purpose, procedures, and data handling and provided consent. The experiment used a three-condition, as shown in Table 1, within-participant repeated-measures design. Participants were divided into six groups of five, and the six possible condition orders were assigned equally to reduce order effects. Each condition lasted 12 days and was followed by a one-day interval before the next condition.

Table 1: Experimental conditions.

Condition	Task	Avatar	Verbal Gratitude
Task-Only	Sudoku	Absent	Absent
Avatar-Response	Sudoku	Visual response	Absent
Avatar-Gratitude	Sudoku	Visual response	After completion

Procedure and Measures

Participants used each application on their own smartphones at times of their choosing. Depressive symptoms and prosocial behavioral intentions were measured at baseline, Day 6, and Day 12. The intention measure comprised 20 concrete helping behavior items adapted for this study. The instructions asked about current intentions on a four-point scale from 'do not want to do' to 'strongly want to do,' and item scores were summed. The concrete-behavior approach resembled established altruism measures (Rushton et al., 1981); however, the present instrument was not a translation of that scale. Therefore, its adaptation from reported behavior to current intention should be considered when interpreting construct validity.

Affective states were measured at baseline, condition end, and after task completion using a Japanese PANAS based on Watson et al. (1988), and a 40-item Japanese Multiple Mood Scale comprising eight five-item subscales. Ishimura and Kodama (2009) provided an English description of an abbreviated application of this method. Depressive symptoms were measured using the Japanese-language CES-D based on Radloff (1977), using a six-day rather than a seven-day recall period to match the measurement interval. After both avatar conditions, participants rated attachment ('I became more attached to the avatar than before use') and helping attribution ('I performed the task for the avatar') on a seven-point scale. The system recorded the total use frequency, number of use days, and maximum consecutive use days.

Statistical Analysis

The primary outcome was the baseline-to-Day 12 change in prosocial behavioral intention, calculated separately for each condition as the Day 12 score minus the baseline score. Because the study used repeated measurements and the change scores were not assumed to be normally distributed, the Friedman test was used to examine whether the changes differed among the three conditions. Pairwise differences between the conditions were examined using Wilcoxon signed-rank tests. The three pairwise p-values were adjusted using Holm's method.

We also examined whether prosocial behavioral intentions changed significantly within each condition. In this analysis, a change score of zero represented no change from the baseline. A one-sample Wilcoxon signed-rank test was conducted for each condition to determine whether the change scores differed from zero, and the resulting three p-values were adjusted using Holm's method.

For Avatar-Gratitude, temporal changes were explored by defining the change from baseline to Day 6 as the early-period change and the change from Day 6 to Day 12 as the late-period change. Cohen's *d* was reported as the effect size for paired comparisons. The significance level was set at 5%.

RESULTS

Prosocial Behavioral Intention

The mean change in prosocial behavioral intention from baseline to Day 12 was −1.467 (SD = 7.921) in the Task-Only, 1.700 (SD = 8.259) in the Avatar-Response, and 3.633 (SD = 6.094) in the Avatar-Gratitude conditions (Figure 3 and Table 2). The omnibus Friedman test did not reach significance (χ^2 (2) = 4.829, p =.089). However, the prespecified Avatar-Gratitude versus Task-Only comparison remained significant after the Holm correction (adjusted p = .042, Cohen's d = 0.53), while the other pairwise comparisons were not significant.

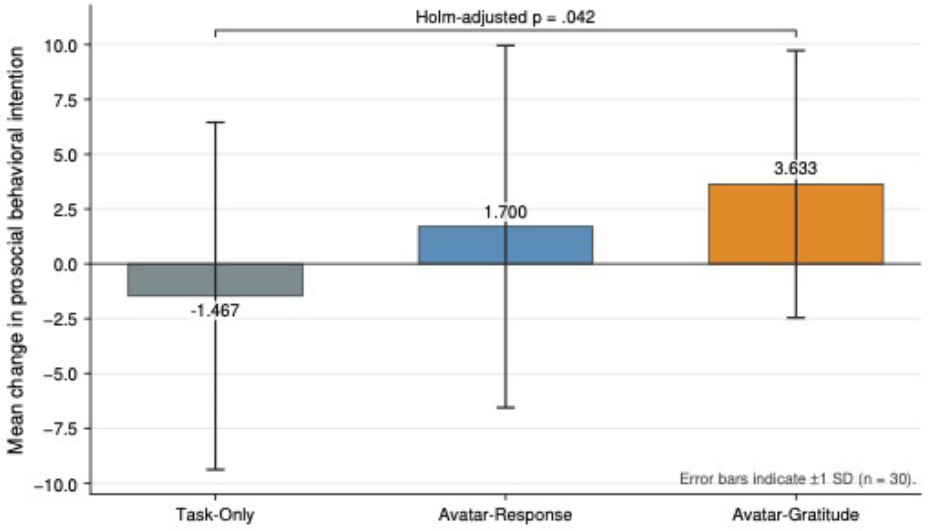


Figure 3: Mean baseline-to-Day 12 changes in prosocial behavioral intention. Error bars show ±1 SD; the *p*-value is Holm-adjusted for the task-only versus avatar-gratitude within-participant comparison.

Table 2: Descriptive and inferential results for prosocial behavioral intention.

Analysis	Mean Change (SD) or Effect Size	Statistical Conclusion
Overall difference among the three conditions	-	Not statistically significant: Friedman $\chi^2(2) = 4.829$; $p = .089$
Task-Only: Baseline to Day 12	−1.467 (SD = 7.921)	No significant change from baseline (adjusted $p > .05$)
Avatar-Response: Baseline to Day 12	1.700 (SD = 8.259)	No significant change from baseline (adjusted $p > .05$)

(Continued)

Table 2: Continued.

Analysis	Mean Change (SD) or Effect Size	Statistical Conclusion
Avatar-Gratitude: Baseline to Day 12	3.633 (SD = 6.094)	Significant increase from baseline (adjusted $p = .015$)
Avatar-Gratitude vs. Task-Only	Cohen's $d = 0.53$	Avatar-Gratitude > Task-Only (adjusted $p = .042$)
Avatar-Gratitude: Baseline to Day 6 (early-period)	2.567 (SD = 5.463)	Significant increase from baseline (adjusted $p = .020$)

Note. $n = 30$. $n = 30$. Values for condition-specific changes are expressed as means (SDs). Change scores were calculated as the later score minus the earlier score; therefore, a zero indicates no change, a positive value indicates an increase, and a negative value indicates a decrease. The omnibus Friedman *p-value* is unadjusted; all other reported *p*-values are Holm-adjusted. Cohen's *d* represents the effect size for the paired comparison.

Affect, Depressive Symptoms, Attachment, and Use

The mean changes in depressive symptoms from baseline to Day 12 were 1.267 in Task-Only, –1.600 in Avatar-Response, and –2.433 in Avatar-Gratitude. The two avatar conditions showed descriptive changes in the favorable direction, but the omnibus difference was not significant ($p = .206$). PANAS and Multiple Mood Scale scores also showed no significant condition differences; the smallest MMS *p*-value for the affiliative affect remained non-significant ($p = .082$).

The mean attachment ratings were 3.600 for Avatar-Response and 4.133 for Avatar-Gratitude ($p = .236$). The ratings for task performance for the avatar were 2.433 and 2.800, respectively ($p = .135$). The mean use frequencies were 3.967, 3.933, and 4.200 for Task-Only, Avatar-Response, and Avatar-Gratitude. Five participants never used Task-Only, five never used Avatar-Response, and two never used Avatar-Gratitude; no usage measure differed significantly among the conditions.

DISCUSSION

Combined Effects of Avatar Responses and Gratitude

The principal finding was that prosocial behavioral intention increased more in Avatar-Gratitude than in Task-Only. Framing a task as help for a virtual entity and returning a positive social response may make users' contributions more visible and meaningful. This interpretation is consistent with the evidence that gratitude communicates social value and motivates subsequent help (Bartlett and DeSteno, 2006; Grant and Gino, 2010).

This finding requires qualification. The omnibus comparison across the three conditions was nonsignificant, as was the comparison between Avatar-Response and Avatar-Gratitude. Therefore, the results only support the possibility that the complete design, which included an avatar, a helping context, a visual response, and verbal gratitude, was more effective than presenting the task alone. They did not establish an independent effect of adding verbal gratitude. A factorial experiment that separates the avatar's

presence, helping instructions, visual response, and verbal gratitude is needed to identify which components contributed to the observed difference.

Feedback Design During Initial Use

The increase in prosocial behavioral intention under Avatar-Gratitude was observed mainly during the first six days. Because the direct comparison between early- and late-period changes was not significant, this pattern did not demonstrate habituation or temporal attenuation. Participants used each application only about four times on average during the 12-day condition, indicating that many did not experience enough of the 15-level affinity progression to reach the plateau. The gradual growth design, which assumed long-term accumulation, was poorly matched to the observed frequency of use.

This result suggests that avatars intended to support prosociality should provide an immediately perceptible response to each helping action from the first use, in addition to more significant changes after long-term engagement. Possible approaches include referring to the user's most recent action, making changes in expression or movement more salient, adapting gratitude messages to use history, and adjusting affinity thresholds to the individual frequency of use. Because usage itself did not differ significantly between conditions, these proposals should be treated as design hypotheses for future evaluation rather than as demonstrated effects.

Effects on Affect and Well-Being

No significant condition differences were found for PANAS scores, Multiple Mood Scale scores, or depressive symptoms. The increase in prosocial behavioral intention therefore cannot be explained by a demonstrated improvement in short-term positive affect or psychological well-being. Although depressive symptoms moved in a favorable direction in both avatar conditions, this was only descriptive. Claims that the system improves well-being will require higher engagement, longer follow-up, and outcome measures designed specifically to assess psychological well-being of the user.

Limitations and Future Work

This study had six main limitations. First, the sample comprised 30 university students, predominantly males. Second, the primary outcome was self-reported intention rather than observed help. Third, the optional use produced substantial variations in exposure. Fourth, the one-day interval may not eliminate carryover. Fifth, Task-Only and Avatar-Gratitude differed in multiple components, preventing the isolation of the independent effect of verbal gratitude. Finally, the secondary and temporal analyses were exploratory, and the study was not powered to establish equivalence when the differences were nonsignificant.

CONCLUSION

This study framed Sudoku completion as helping an avatar and evaluated visual responses and gratitude feedback in a 12-day, three-condition repeated measures experiment. Prosocial behavioral intention increased more in Avatar-Gratitude than in Task-Only, but the omnibus comparison and the Avatar-Gratitude versus Avatar-Response contrast were nonsignificant. Therefore, the evidence concerns the combined avatar feedback experience, not the independent causal effect of verbal gratitude. No reliable improvements in affect, depressive symptoms, attachment, or usage were observed. The low frequency of use further indicates that socially responsive avatars should provide immediate and salient feedback and adapt long-term progression to actual engagement. These findings offer a bounded but practical basis for designing and testing avatar-mediated interventions that support prosocial behavior.

ACKNOWLEDGMENT

This study was supported by JSPS KAKENHI (grant no. 23K11294). This study builds on Kenta Sekiuchi's undergraduate thesis and extends the original work through additional analyses.

REFERENCES

Bartlett, M. Y. and DeSteno, D. (2006). "Gratitude and prosocial behavior: Helping when it costs you", Psychological Science, 17(4), 319–325. https://doi.org/10.1111/j.1467-9280.2006.01705.x

Baylor, A. L. (2009). "Promoting motivation with virtual agents and avatars: Role of visual presence and appearance", Philosophical Transactions of the Royal Society B: Biological Sciences, 364(1535), 3559–3565. https://doi.org/10.1098/rstb.2009.0148

Calvo, R. A., and Peters, D. (2014). Positive Computing: Technology for Wellbeing and Human Potential. MIT Press. https://doi.org/10.7551/mitpress/9764.001.0001

Grant, A. M., and Gino, F. (2010). "A little thanks goes a long way: Explaining why gratitude expressions motivate prosocial behavior", Journal of Personality and Social Psychology, 98(6), 946–955. https://doi.org/10.1037/a0017935

Hu, T.-Y., Li, J., Jia, H. and Xie, X. (2016). "Helping others, warming yourself: Altruistic behaviors increase warmth feelings of the ambient environment", Frontiers in Psychology, 7, 1349. https://doi.org/10.3389/fpsyg.2016.01349

Ishimura, I. and Kodama, M. (2009). "Flow experiences in everyday activities of Japanese college students: Autotelic people and time management", Japanese Psychological Research, 51(1), 47–54. https://doi.org/10.1111/j.1468-5884.2009.00387.x

Jahn, K., Kordyaka, B., Machulska, A., Eiler, T. J., Gruenewald, A., Klucken, T., Brueck, R., Gethmann, C. F. and Niehaves, B. (2021). "Individualized gamification elements: The impact of avatar and feedback design on reuse intention", Computers in Human Behavior, 119, 106702. https://doi.org/10.1016/j.chb.2021.106702

Radloff, L. S. (1977). "The CES-D Scale: A self-report depression scale for research in the general population", Applied Psychological Measurement, 1(3), 385–401. https://doi.org/10.1177/014662167700100306

Rushton, J. P., Chrisjohn, R. D. and Fekken, G. C. (1981). "The altruistic personality and the Self-Report Altruism Scale", Personality and Individual Differences, 2(4), 293–302. https://doi.org/10.1016/0191-8869(81)90084-2

Watson, D., Clark, L. A. and Tellegen, A. (1988). "Development and validation of brief measures of positive and negative affect: The PANAS scales", Journal of Personality and Social Psychology, 54(6), 1063–1070. https://doi.org/10.1037/0022-3514.54.6.1063