Advances in Mental Health Care: Five N = 1 Studies on the Effects of the Robot Seal Paro in Adults With Severe Intellectual Disabilities

Eline Wagemaker, Tycho J. Dekkers, Joost A. Agelink van Rentergem, Karin M. Volkers & Hilde M. Huizenga

To cite this article: Eline Wagemaker, Tycho J. Dekkers, Joost A. Agelink van Rentergem, Karin M. Volkers & Hilde M. Huizenga (2017) Advances in Mental Health Care: Five N = 1 Studies on the Effects of the Robot Seal Paro in Adults With Severe Intellectual Disabilities, Journal of Mental Health Research in Intellectual Disabilities, 10:4, 309-320, DOI: 10.1080/19315864.2017.1320601

To link to this article: https://doi.org/10.1080/19315864.2017.1320601
Advances in Mental Health Care: Five $N=1$ Studies on the Effects of the Robot Seal Paro in Adults With Severe Intellectual Disabilities

Eline Wagemaker$^{a,b}$, Tycho J. Dekkers$^{a,c}$, Joost A. Agelink van Rentergem$^{a}$, Karin M. Volkers$^{b}$, and Hilde M. Huizenga$^{a,d,e}$

$^a$Department of Psychology, University of Amsterdam, Amsterdam, the Netherlands; $^b$Philadelphia Care Foundation, Amersfoort, the Netherlands; $^c$Department of Forensic Youth Psychiatry and Complex Behavioral Disorders, De Bascule, Academic Center of Child and Adolescent Psychiatry, Duivendrecht, the Netherlands; $^d$Research Priority Area Yield, University of Amsterdam, Amsterdam, the Netherlands; $^e$Amsterdam Brain and Cognition Center, University of Amsterdam, Amsterdam, the Netherlands

ABSTRACT

Background: The evidence base for psychological treatments for autism and mood disorders in people with moderate to severe intellectual disabilities (ID) is limited. Recent promising robot-based innovations in mental health care suggest that robot-based animal assisted therapy (AAT) could be useful to improve social skills and mood in people with moderate to severe ID. This study explores the efficacy of the innovative therapeutic robot seal Paro on alertness and mood in adults with moderate to severe ID. Method: Five $N=1$ studies were performed in adults (59–70 years) with moderate to severe ID. During a control phase of four weeks, participants could interact with a plush seal, which was replaced by the robot seal Paro during an equally long treatment phase. Results: In one participant, alertness and mood, as rated by supervisors, improved during the treatment phase as compared to the baseline phase. The other four participants did not improve on either of the outcome measures. Conclusion: We conclude that robot-based AAT does not have clear beneficial effects on alertness and mood in adults with moderate to severe ID, but that positive interactions with the robot seal could be of therapeutic value in itself.

Background

Psychopathology is highly prevalent in people with severe intellectual disabilities (ID). For example, the point prevalence in a sample of people with moderate to severe ID was 8.6% for mood disorders and 5.6% for autism (Cooper, Smiley, Morrison, Williamson, & Allan, 2007). However, the evidence base for psychological treatments to improve the abovementioned disorders in this group is limited (Bhaumik, Gangadharan, Hiremath, & Russell, 2011). More knowledge about the efficacy of treatments in people with moderate to severe ID is necessary to improve...
their quality of life. Therefore, the current study is focused on an innovative treatment for social and emotional symptoms in adults with moderate to severe ID.

A recent innovation in many mental healthcare institutions is the application of socially assistive robots. These are employed for a variety of purposes in all kinds of groups. For example, robots are employed as a therapeutic play partner for children with autism, as a coach to provide direct instruction and supervision to elderly people engaged in memory exercises, or as a companion to prevent loneliness in the elderly with dementia (see Rabbitt, Kazdin, & Scassellati, 2015, for a review).

The most investigated socially assistive robot is the robot seal Paro (Wada, Shibata, Saito, & Tanie, 2004). Paro has been developed to provoke social reactions by having soft fur, by reacting to its environment with sounds, and by moving its head and tail. The first results on the use of Paro in the elderly (with and without dementia) are promising: A review showed positive effects of Paro on social interaction and mood (Bemelmans, Gelderblom, Jonker, & De Witte, 2012). For example, demented elderly people improved with regard to social skills and stress reactivity (Wada & Shibata, 2007) and mood and communication (Wada, Shibata, Musha, & Kimura, 2008) after interaction with Paro. More recent randomized controlled trials (RCTs) suggest that Paro is more effective in producing these social and emotional benefits than other treatments. For example, an RCT in a group of elderly people without dementia reported that loneliness decreased significantly in the Paro group as compared to the control group that interacted with a resident dog (Robinson, MacDonald, Kerse, & Broadbent, 2013). Relatedly, studies using AB designs showed more interest in Paro than in a stuffed lion in demented elderly people and more social interaction and play for the elderly without dementia when Paro was turned on (Takayanagi, Kirita, & Shibata, 2014; Kidd, Taggart, & Turkle, 2006). To conclude, the evidence for the efficacy of Paro in improving social and emotional aspects in elderly people is growing. Potentially, Paro would have the same beneficial effects in adults with moderate to severe ID.

Relatedly, animal-assisted therapy (AAT) is frequently used in mental health care institutions for individuals with ID (Fine, 2010). Studies on AAT in ID show mixed results. Beneficial effects during animal interaction have been reported (e.g., Esteves & Stokes, 2008), but the advantage of AAT compared to control conditions and the generalization outside the treatment setting are questionable (e.g., Fung & Leung, 2014 but see Griffioen & Enders-Slegers, 2014). Moreover, some practical issues arise when using animals within mental health-care institutions, such as caretaking, cleaning, allergies and safety issues. Thus, it may be worthwhile to consider modifications.

Given the aforementioned robot-based innovations in mental health care, and the potentially effective application of AAT in ID, the current study investigates the
potential of robot-based AAT in adults with moderate to severe ID using the robot seal Paro. Robot-based AAT could be a good alternative for AAT. The practical issues mentioned above do not arise, as robots do not need any further care, are easy to clean, do not give rise to allergies, and are safe. As a result, robot-based AAT does not put an additional burden on the high workload of the staff in mental health-care institutions (Skirrow & Hatton, 2007). If the efficacy of a robot is proven, it could be implemented without much effort.

The aim of the current study was to explore the efficacy of Paro in adults with moderate to severe ID on mood and alertness. As mentioned before, mood disorders are not uncommon in adults with moderate to severe ID (Cooper et al., 2007) and Paro was able to improve mood in the elderly (Bemelmans et al., 2012). Furthermore, reductions and irregular changes in alertness were reported in adults with moderate to severe ID (Munde, Vlaskamp, Vos, Maes, & Ruijsenaars, 2012). Increasing alertness would be beneficial, as alertness is an important precondition for both learning and communication (Guess, Roberts, & Guy, 1999). This could especially be helpful for adults with ID and autism. Moreover, simple interactions with Paro may match the low cognitive functioning of adults with moderate to severe ID. To the authors’ knowledge, only one study investigated the use of Paro in people with ID. This study described an increase in activity in a child with severe ID, as well as improvements in communication in a twin with Angelmann syndrome after interacting with Paro (Marti et al., 2005). Therefore, Paro seems a promising intervention to improve mood and alertness in adults with moderate to severe ID.

Since the use of Paro in ID is relatively new, the current study uses an $N=1$ design instead of an RCT. $N=1$ studies can provide first clues about potential efficacy of a therapy and are less expensive than large RCTs (Barlow, Nock, & Hersen, 2009; Kazdin, 1978, 2011; Maric, De Haan, Hogendoorn, Wolters, & Huizenga, 2015). The $N=1$ study consists of an active control phase and an equally long treatment phase (i.e., AB design; Barlow et al., 2009). On mood and alertness, we expect (1) improvement of outcomes over time during the treatment phase, (2) more pronounced improvements during the treatment than during the control phase and, (3) better ratings at the end of treatment than at the end of the control phase.

**Method**

**Participants**

Five participants (59–70 years) with a moderate (IQ < 50) to severe (IQ < 35) ID, but without symptoms of dementia, were included. All participants were living in the same group of a residential mental health-care institution (see Table 1 for background characteristics). Informed consent was given by the legal representative of the participants.
**Procedure**

The study was approved by the institutional review board. It consisted of two phases, both with a duration of four weeks. In the control phase, a plush seal (named Tobi) with a very similar appearance to the robot seal (Paro) was present in the living room of the participants. In the treatment phase, the plush seal was replaced by the robot seal. Supervisors received a structured protocol on how to present Tobi and Paro to the participants.

On the first day of both phases, the seal was explicitly introduced during breakfast, and all participants were asked to touch the seal. During the first week of both phases, in order to get participants’ attention, in the morning the seal was explicitly greeted “good morning” by the supervisors and was put on the table during breakfast every day. Every evening, the seal was put into every participants’ lap for two minutes, and during these two minutes the supervisor stayed close and modeled some interaction with the seal. After two minutes, the supervisor withdrew and the participant self-determined the duration of interaction with the seal. If participants reacted anxiously or rejecting, which happened occasionally, supervisors stopped the interaction. At the end of the evening, the supervisor and participants said goodnight to the seal together.

During the other weeks of both phases, a similar ritual was used in the morning and in the evening. However, the breakfast ritual and the two-minutes guided interaction with the seal decreased from seven to three days a week. In the evening, the seal was put on the lap of those participants who approved to interact with the seal. The participant determined the duration of this interaction. During all weeks, participants were free to interact with the seal as much as desired when they were at home, because the seal was constantly available in the living room.

During eight weeks, mood and alertness ratings were provided twice every day by the daily supervisors (morning and evening shift) and self-reported mood ratings were provided once a day (before bedtime). Additionally, supervisors were asked for qualitative observations on how each participant reacted to the seal throughout the study.

<table>
<thead>
<tr>
<th>Participant</th>
<th>Gender</th>
<th>Age</th>
<th>ID</th>
<th>Characteristics</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>M</td>
<td>63</td>
<td>SID</td>
<td>Fragile X syndrome, autism spectrum disorder, epilepsy, history of depressive episode</td>
</tr>
<tr>
<td>2</td>
<td>F</td>
<td>70</td>
<td>SID</td>
<td>History of depressive and psychotic episodes</td>
</tr>
<tr>
<td>3</td>
<td>F</td>
<td>60</td>
<td>SID</td>
<td>Autism spectrum disorder, history of psychotic episode</td>
</tr>
<tr>
<td>4</td>
<td>F</td>
<td>59</td>
<td>SID</td>
<td>Epilepsy, history of depressive episode</td>
</tr>
<tr>
<td>5</td>
<td>M</td>
<td>68</td>
<td>MID/SID</td>
<td>Epilepsy</td>
</tr>
</tbody>
</table>

*Note. M = male, F = female, SID/MID = severe/moderate intellectual disability.*
**Materials**

**Seals**

Paro is a robot seal that is equipped with multiple sensors, processors, and microphones, which altogether gives a realistic impression. The robot can move in different ways, its eyes can open and close automatically, and it responds to touching and to specific sounds (see Wada et al., 2004 for a more extensive description of Paro). The plush seal that is used during the control phase has a very similar appearance, although it is somewhat smaller and has no interactive features.

**Mood**

Existing mood questionnaires were inappropriate, as participants were not able to read and write sufficiently. We therefore asked participants to point at one out of three smileys that best described their mood at that moment. The smileys were sad/angry, neutral, or happy, similar to the Young Child Session Rating Scale (Duncan, Miller, & Sparks, 2011). Mood estimations were given every day by means of self-report, but were also provided on the same scale by the daily supervisors twice (morning and evening shift), altogether summing to 56 self-report measures and 112 measures by the daily supervisor for every participant.

**Alertness**

Alertness of the participants was scored on the Alertness Observation List (AOL; Vlaskamp, Fonteine, Tadema, & Munde, 2010) twice every day by the daily supervisor (morning and evening shift). Supervisors had to divide 100% over a profile of alertness that consisted of four categories: active and directed at the environment, inactive and withdrawn, sleepy/drowsy, agitated (Munde et al., 2012). A reliability study on the AOL in adults with profound ID found an interobserver reliability of $r = .81$ and an intra-observer reliability of $r = .87$ (Munde et al., 2011), suggesting that the AOL is a reliable instrument. In total, alertness was measured at 112 occasions for every participant.

**Data Analysis**

An $N = 1$ regression analysis was performed for every participant, correcting for time-dependent residuals (Maric et al., 2015). The regression model was described by the following equation:

$$y(i) = b_0 + b_1 \times \text{phase}(i) + b_2 \times \text{time.in.phase}(i) + b_3 \times \text{phase}(i) \times \text{time.in.phase}(i) + e(i)$$

(1)
In this equation, \( y(i) \) denotes the dependent variable score at time point \( i \), \( \text{phase}(i) \) denotes the phase in which the time point \( i \) is contained (coded as 1 for baseline and 0 for treatment) and \( \text{time.in.phase} \) denotes time points within each phase. The term \( e(i) \) denotes the residual at time point \( i \). To compare differences between phases as well as differences within phase, \( \text{time.in.phase} \) is coded in reverse order so the last time point of each phase equals 0. Consequently, in the model the parameter \( b_0 \) denotes the value of the dependent variable at the end of the treatment, \( b_1 \) denotes the change in value between the end of the baseline and the end of the treatment, \( b_2 \) denotes the time effect on the dependent variable in the treatment phase, and \( b_3 \) denotes the change in time effect on the dependent variable in the baseline phase compared to the time effect in the treatment phase.

By using this model, both differences between phases as well as differences within phases can be estimated. The model holds under the following assumptions: effect of time in both phases is linear, residuals are normally distributed, and residuals are correlated according to an auto-regressive structure (Maric et al., 2015). Analyses were performed with the freely available E-CLIP N = 1 app (Agelink Van Rentergem & Huizenga, 2016).

**Results**

**Missing Data**

The overall percentage of missing data was 17.6%. For mood as rated by the supervisor, an average of 94.4 (out of a possible 112) measurements per participant were obtained. For self-rated mood, an average of 43.6 (out of a possible 56) measurements were obtained. For alertness, an average of 95.2 (out of a possible 112) measurements per participant were obtained. Missing data were likely to be the result of the high demand of the study on daily supervisors, who sometimes forgot to fill in the ratings, especially at the end of the control phase. However, given the large amount of acquired data, it is unlikely that these missing data influenced the results.

**Participant 1**

Time had a positive effect on self-reported mood during treatment (\( b = -0.03, t = -2.16, p < .05 \); see Figure 1). However, the time effect during treatment was not different from the effect of time during the control phase (\( b = 0.02, t = 0.97, p = .34 \)), and mood at the end of treatment did not differ from mood at the end of the control phase (\( b = -0.26, t = -0.78, p = .44 \)). For this participant, no significant effects on mood as well as alertness as rated by the daily supervisors were found.
Participant 2
Surprisingly, in this participant, mood ratings by the supervisors tended to be higher at the end of the control phase than at the end of the treatment phase ($b = 0.35$, $t = 1.80$, $p = .08$). No significant effects on self-reported mood or on alertness as rated by supervisors were found.

Participant 3
In participant 3, self-rated mood at the end of the treatment tended to be better than mood at the end of control phase ($b = -0.60$, $t = -1.78$, $p = .08$). No other effects were found on the self-report measure of mood. With regard to mood as rated by daily supervisors (see Figure 2), higher scores were observed at the end of the treatment phase as compared to the end of the control phase ($b = -0.59$, $t = -2.18$, $p < .05$). In the same participant, there was a positive effect of time on mood ratings during the treatment phase ($b = -0.01$, $t = -2.11$, $p < .05$) and this effect of time was higher in the treatment phase than in the control phase ($b = 0.02$, $t = 2.17$, $p < .05$).

With regard to the alertness profile, an increase in activity and directedness at the environment during the treatment phase was observed ($b = -0.44$, $t = -1.99$, $p < .05$; see Figure 3a). In addition, activity and directedness at the environment was higher at the end of the treatment phase than at the end of the control phase ($b = -31.53$, $t = -2.96$, $p < .01$). However, the effect of time did not differ between both phases ($b = 0.45$, $t = 1.36$, $p = .18$). Furthermore, lower agitation was observed at the end of the treatment phase as compared to the end of the control phase ($b = 16.45$, $t = 2.35$, $p < .05$; see Figure 3b). Although there was a trend toward a difference in the effect of time between both phases ($b = -0.37$, $t = -1.72$, $p = .09$), the effect of time in the treatment phase was not significant ($b = 0.12$, $t = 0.85$, $p = .40$). No significant effects were observed on the other categories of the alertness profile.
Figure 2. Mood ratings of participant 3, as rated by the daily supervisors.

Figure 3. Alertness ratings by daily supervisors of participant 3: (a) activity and directedness at environment, (b) agitation.
**Participant 4**
In this participant, no effects were found on any outcome measure, except a trend for a decrease in agitation during the treatment phase ($b = 0.12$, $t = 1.94$, $p = .06$).

**Participant 5**
No significant effects were found for this participant.

**Informal Observations**
Informal observations of the daily supervisors of the participants indicated that the robot seal attracted more attention than the plush seal. The supervisors were enthusiastic about the robot, and regarded its presence as positive. At the introduction of the robot, mild anxiety was observed in some participants, presumably because of the unpredictable sounds and movements of the robot, but this diminished quickly. Supervisors reported that they were of the opinion that Paro was especially useful for participant 2. She had long and positive interactions with the seal and seemed to establish an emotional connection.

**Conclusion**
The aim of this study was to explore the efficacy of the innovative robot seal Paro in five adults with moderate to severe ID by extensive observation of the adults. In one participant, in line with expectations, Paro had a positive influence on mood and alertness as rated by the daily supervisors. In the other four participants, contrary to expectations, no significant beneficial effects of Paro were observed. Therefore, the current results are not very promising with regard to the efficacy of the robot seal Paro in older adults with moderate to severe intellectual disability.

This study has several limitations. Most importantly, the current study investigated only five participants, which precludes generalization to other individuals. However, such $N = 1$ studies are essential as they can provide first clues about potential efficacy of therapy and are less expensive than large randomized control trials (Barlow et al., 2009; Kazdin, 1978; Maric et al., 2015). A second limitation is that an ABAB design consisting of four phases instead of the current AB design was impossible for practical reasons and because of ethical concerns from the mental health care institution. An ABAB design might have increased reliability of findings because it enables the distinction between the effects of time from the effects of phase. On the contrary, in the AB design improvements from the A to the B phase could have been merely due to time instead of phase. However, it seems unlikely that an ABAB design would have yielded different results, as for most participants, no differences were observed between the two phases. Third,
the self-report scale on mood might have been too complicated for some of the participants, as was reported afterward by the daily supervisors.

More generally, the potential reduction in human contact is the most often mentioned ethical concern about the use of robots like Paro in mental health care (Sharkey & Sharkey, 2012). By all means, it is not desirable to replace human contact by robot contact. Therefore, robot-based care is at most an addition to the already intensive human contact in the care for people with severe intellectual disabilities. To guarantee ethical application in clinical practice, mental health care institutions should regularly evaluate if contact with Paro does not imply decreased human contact with caretakers.

Informal observations of daily supervisors indicated the most positive effects of the robot seal in participant 2, who interacted most frequently with Paro, seemed to establish an emotional connection, and showed mood improvements during the interaction. However, this participant did not improve on the chosen outcome measures. It might well be the case that generalization of treatment effects in this group of severely intellectually disabled adults is an unrealistic aim. Positive interaction with the animals might be an appropriate goal in itself (see, e.g., Esteves & Stokes, 2008), even when this interaction does not lead to sustainable long-term effects.

To conclude, this study shows that it is feasible to implement and investigate robot-based AAT in residential mental health care institutions for moderately to severely intellectually disabled adults. Based on the results of five extensive \( N = 1 \) studies, we recommend that mental health care institutions be cautious with purchasing expensive therapeutic robot seals, as the efficacy in comparison to a plush seal is not evident. However, it must be noted that the clinical impression of supervisors regarding the robot seal was generally positive, and that in this population positive interaction with an animal-like robot might be a therapeutic aim in itself.

**Acknowledgments**

The authors are very thankful to the five participants and their caretakers for their participation in this study. Their gratitude also goes to Marianne Swart and all the daily supervisors at Philadelphia Care Foundation, for their enthusiasm in collaborating in this study and for collecting all the data.

**References**


