An exploratory study into the use of an emotionally aware cognitive assistant

Aarti Malhotra [1], Lifei Yu [1], Tobias Schröder [2], Jesse Hoey [1]
[1] School of Computer Science, University of Waterloo, Waterloo, Ontario, N2L3G1
[2] Potsdam University of Applied Sciences, 14469 Potsdam, Germany
{aarti.malhotra, l38yu, jhoey}@uwaterloo.ca, post@tobiasschroeder.de

Abstract

This paper presents an exploratory study conducted to understand how audio-visual prompts are understood by people on an emotional level as a first step towards the more challenging task of designing emotionally aligned prompts for persons with cognitive disabilities such as Alzheimer's disease and related dementias (ADRD). Persons with ADRD often need assistance from a caregiver to complete daily living activities such as washing hands, making food, or getting dressed. Artificially intelligent systems have been developed that can assist in such situations. This paper presents a set of prompt videos of a virtual human 'Rachel', wherein she expressively communicates prompts at each step of a simple hand washing task, with various human-like emotions and behaviors. A user study was conducted for 30 such videos with respect to three basic and important dimensions of emotional experience: evaluation, potency, and activity. The results show that, while people generally agree on the evaluation (valence: good/bad) of a prompt, consensus about power and activity is not as socially homogeneous. Our long term aim is to enhance such systems by delivering automated prompts that are emotionally aligned with individuals in order to help with prompt adherence and with long-term adoption.

1. Introduction

People with cognitive disabilities such as Alzheimer's disease and related dementias (ADRD) have trouble completing activities of daily living (AD) and are usually assisted by a human care partner. The use of computerized intelligent cognitive assistants (ICAs) can help reduce a care partner's burden and help a person suffering from ADRD to perform the tasks more independently. This also increases independence and control and thereby reduces depression and feelings of powerlessness and dependence.

Copyright © 2015, Association for the Advancement of Artificial Intelligence (www.aaai.org). All rights reserved.

These ICAs take the form of automated methods for monitoring a person and inferring their activities and needs, combined with some form of prompting to provide them assistance when necessary. However, even when this technology satisfies functional requirements, people often reject it. We believe that a major reason for non-adoption is the lack of an affective (emotional) connection between technology and human.

For example, (Mihailidis et al. 2008) demonstrates a prompting system called the COACH that can monitor a person with Alzheimer's disease while they are trying to wash their hands, detect when they have lost track of what they are doing, and play a prerecorded assistive prompt. The COACH is effective at monitoring and making decisions about when/what to prompt (Mihailidis et al. 2008), and works well for some persons, but not as well for others. Considering the heterogeneity in socio-cultural and personal affective identities, a primary reason for lack of effectiveness may be the static, non-adaptive nature of the "canned" (pre-recorded) prompts. While we have made significant effort to design prompts founded on the methods and styles of human caregivers (Wilson et al. 2013), a simple "one size fits all" style of prompting may be limiting. While one person may find a prompt helpful and motivational, another may find it imperious and impatient. The first person is likely to follow the prompt, to feel respected, valued, and utilize the technology. The second person, on the other hand, may feel confused by the prompt and discontinue the task. However, a different style of prompting (i.e. a more subtle prompt, perhaps with a different tone of voice, or with a different wording), may be much more effective for the second person, but not for the first. Each person comes from a different background, has a different sense of "self", and has different emotional responses to prompts. Affective identity is believed to be a powerful tool for reasoning about illness in general (Lively and Smith 2011). Studies of identity in Alzheimer's disease have found that identity changes dramatically over the course of the disease (Orona 1990), and that persons with ADRD have more vague or abstract notions of their identity (Rose Addis and Tippett 2004).

Our long-term aim is to build technology that will detect and adapt to these differences. In this paper, we report upon a first step towards this aim by creating a set of audio-visual prompts using a virtual human developed with the USC Virtual Human Toolkit (VHT). We built a set of six such prompts (for different steps of the handwashing task) with five different personalities (e.g. "bossy", "motherly" or "bored"). We then did an online survey to measure human responses to these 30 different prompts on three important emotional dimensions of Evaluation (valence), Potency (power) and Activity, termed as EPA. We present results from 27 respondents to this survey and analysis of the results primarily in terms of the *consensus* of respondents within each dimension (EPA). We show that while the respondents tend to agree on the *evaluation* dimension ("good" vs. "bad"), there is less clear agreement on the potency ("powerful" vs. "powerless") and activity ("active" vs. "asleep") dimensions. These considerations will be important in the pursuit of widely accepted and personally effective assistive technologies.

In order to develop prompts for assistive technologies, we will need to do a survey such as this with participants strictly from the target user group (elder persons with cognitive disabilities). However, this is not possible at such an early stage, due to the challenges posed by this population and the lack of any previously published work on automated emotionally aligned prompts. Without prior work, it would be very difficult to get ethical approval for studying prompts that react to and change the emotional state of persons with Alzheimer's disease, who generally cannot provide informed consent. Thus, surveying non-cognitively disabled persons is a critical building block towards an eventual survey of the target user group.

2. Background

The task of hand washing activity consists of five essential steps: turning water on, using soap, rinsing hands, drying hands, and turning water off. An assistive handwashing system called COACH (Cognitive Orthosis for Assisting with aCtivites in the Home) (Mihailidis et al. 2008) uses a video camera placed above the wash basin which captures the current activity by tracking hand and towel positions. An artificial intelligence module determines an appropriate action to take: either 'prompt the user' for one of steps of handwashing, or 'summon the caregiver' or 'continue observing the user activity'. The speaker of the prompt is not currently visualized which limits the interaction.

Affect Control Theory (ACT) arises from work on the sociology of human interaction (Heise 2007). ACT proposes that social perceptions, behaviors, and emotions are guided by a psychological need to minimize the differences between culturally shared fundamental affective sentiments about social situations and the transient impressions resulting from the interactions between elements within those situations. Both the fundamentals and transients are represented as vectors in a three-dimensional affective space, the basis vectors of which are called Evaluation, Potency, and Activity (EPA). The EPA space is hypothesized to be a universal organizing principle of human socio-emotional experience, based on the discovery that these dimensions structure the semantic relations of linguistic concepts across languages and cultures (Osgood, May, and Miron 1975). They also emerged from statistical analyses of the co-occurrence of a large variety of physiological, facial, gestural, and cognitive features of emotional experience (Fontaine et al. 2007), and relate to the universal dimensionality of personality, and social cognition (Scholl 2013).

EPA profiles of concepts can be measured with the semantic differential, a survey technique where respondents rate affective meanings of concepts on numerical scales. In general, within-cultural agreement about EPA meanings of social concepts is high even across subgroups of society, and cultural-average EPA ratings from as little as a few dozen survey participants are extremely stable over extended periods of time (Heise 2010). For example, the EPA for the identity of "nurse" is [1.65,0.93,0.34], meaning that nurses are seen as quite good (E), a bit powerful (P), and a bit active (A)[1]. Comparatively a "patient" is seen as [0.9,−0.69,−1.05], less powerful and less active than a "nurse". Social events cause transient impressions of identities and behaviors that deviate from their corresponding fundamental sentiments. ACT models this formation of impressions from events with a minimalist grammar of the form agent-behavior-client. Consider, for example, a nurse (agent) who ignores (behavior) a patient (client). Observers agree, and ACT predicts, that this nurse appears less nice (E), and less potent (P), than the cultural average of a nurse. The weighted sum of squared Euclidean distance between fundamentals and transients is called the *deflection* and is hypothesized to correspond to an aversive state of mind that humans seek to avoid or minimize (the affect control principle). The nurse who "ignores" a patient has a deflection of over 15 (very high), whereas if the nurse "comforts" the patient, the deflection is 1.5 (very low). The affect control principle also allows ACT to compute normative deflection minimizing actions for artificial agents. ACT has been shown to be a powerful predictor of human behavior (MacKinnnon and Robinson 2014).

[1] EPA values range from -4.3 to 4.3 by convention

Recently, a probabilistic and decision theoretic generalization of the ACT model was proposed called BayesAct (Hoey, Schröder and Alhothali 2013; Hoey and Schröder 2015). BayesAct allows ACT principles to be used to guide artificially intelligent systems on an emotional level. It also allows ACT to model more complex affective sentiments, including multi-modal ones as it considers sentiments as probability distributions rather than just points in EPA space. When used in the enhanced COACH system, this engine chooses a deflection minimizing action according to ACT principles.

In this paper, we develop a set of video prompts involving a virtual human character that can be used by the COACH system enhanced with the BayesAct engine. The measured EPA values given by this study provide a reference for the resulting system to choose the best (emotionally aligned) video prompt for a desired user action step. This would improve system-user interaction at each step of the handwashing task and enrich overall user experience of the system. A proof of concept system using these prompts and integrating COACH with BayesAct was presented in (Lin et al. 2014). Here, we focus on the prompt generation and initial user survey.

3. Video Prompts Design

The videos generated for the prompts consisted of a virtual human character called 'Rachel' created with the VHT[2]. We used the *NonVerbal Behavior Generato*r, the *SmartBody* module, and the *Character Customizer* tools that allow for the quick setup of a single-character scene and a set of lines for the character to act out. These modules offer control over camera angles, backgrounds, voices and facial animation. The prompts were entered and the facial expression was configured for each prompt. Certain behaviors were assigned for certain words, so that the character could display accordingly when those words were spoken. Speech was used from one of the standard options 'Microsoft Anna'. The sub-title option was off.

3.1 Prompts Design

We chose to evaluate 5 archetypal "personalities" for prompting, corresponding to EPAs of: + + +, + - -, - + +, - - -, and + - +. We refer to these as the "Expected EPA" values as given in Table 1. These personalities corresponded roughly to the identities of "big sister", "grandmother", "politician", "bore", and "teenager", respectively, according to the sentiment repository distributed with (Heise 2007). These EPA profiles were selected as those we could represent with the VHT and that spanned the space of usual affective identities.

[2] Virtual Human Toolkit (https://vhtoolkit.ict.usc.edu)

Table 1. Behavior/Identities and Expected EPA

Behavior/identities	**Expected EPA**
discipline/big sister/supervise	+ + +
request/granny/bow to	+ - -
bossy/politician	- + +
unadventurous/bore	- - -
impatient/teenager/little brother	+ - +

We had 5 functional prompts, as in (Mihailidis et al. 2008), and added one more for goodbye. Hence we had 30 videos to be surveyed (6 functional x 5 personalities). Table 2 shows the set of prompts (linguistic component only) for the 'using soap' step of handwashing task. Complete set of prompts is in (Malhotra et al. 2014).

Table 2. Prompts for handwashing step 'using soap'

Behavior/identities	**Prompt**
discipline/big sister/supervise	*Try putting on some soap.*
request/granny/bow to	*You are washing your hands. Please use the soap.*
bossy/politician	*NOW use the soap*
unadventurous/bore	*If you want to put on some soap, there is a soap pump lying around*
impatient/teenager/little brother	*I want you to put on some soap*

Figure1: Example frames from a "Bossy" prompt to dry hands

Based on Expected EPA we devised a set of non-verbal behavior rules in the Virtual Human Toolkit to match with the speech. For example, to have our character 'Rachel' deliver the prompt in a bossy manner, we put in a rule so that whenever the word 'NOW' occurs in the prompt, the character would use an animation, which depicts a commanding behavior by moving a closed fist similar to a beating movement in an up and down fashion, as shown in Figure 1. Also the facial expression was customized to have an angry look on Rachel's face to complement the speech and behavior. The non-verbal rules were designed by the authors according to their intuitive feelings about the expected identities and behaviors as shown in Table 1. The set of video prompts generated here span a reasonable region of the affective space and the user study performed measures mapping of gestures to the affective space.

4. User Study

An online survey was conducted in which participants were asked to watch the 30 videos and rate them based on Evaluation, Potency, and Activity dimensions (on a discrete scale of -4 to +4 with increments of 1 for a total of 9 options). We applied the standard methodology of the semantic differential, as developed and validated by Osgood (described in detail in (Heise 2010). We showed sets of concepts at either end of the scales as follows:

- **Evaluation**: bad/awful to good/nice
- **Potency**:impotent/powerless/little to potent/powerful/big
- **Activity**: inactive/slow/quiet to active/fast/noisy

These combinations of adjectives are meant to reduce the effects of concept-scale interactions (where the words defining the scales cause changes in participants responses on the absolute scale values) (Heise 2010). Further, the use of a single scale for each dimension is a tradeoff of measurement and economy. The survey took 10 minutes to complete, a significant barrier to getting sufficient respondents. We provided meaning of all three dimensions to the participants before they started the survey and put labels for the ranges to guide them.

The questions were presented in randomized order and the survey was kept active for three weeks. The survey was advertised on local and international mailing lists. Participants were shown a setup video before providing their ratings, so they could confirm if their video and audio was working. They could skip any question or exit the survey at any point in time. At the end of the survey, information on gender, age group, and free-form comments were requested. In the end, an appreciation video was shown. There were total of 27 respondents (16 male/9 female with 18 Canadians and 9 internationals) who answered more than 90% of questions. An example screen shot showing one of the questions is shown in Figure 2.

4.1 Consensus Analysis

To determine consensus amongst participants, we followed the culture-as-consensus model measuring the shared knowledge of the culture within the respondents (Borgatti and Halgin 2010; Schröder et al., 2013; Heise 2010). This allows measurement of whether respondents homogeneously represent a common culture that explains the similarity in their answers to questions about the cultural norms (Romney et al. 1986). The method computes the Eigenvalues of the covariance matrix of all responses for each of E,P,A separately, where the covariance is computed in the space spanned by the participants, with data given by the questions. Thus, we are computing the principal components that indicate the extent to which respondents agree in their ratings across all items. If there is one large first factor (i.e. the first Eigenvalue is notably larger than the second Eigenvalue), this reflects cultural commonality in the respondent's ratings and provides evidence of one dominant factor governing respondent's judgement (Heise 2010). In this paper, if the first-to-second Eigenvalue ratio is equal or greater than 2.0, it is considered as significant, as suggested by (Romney et al. 1986; Heise 2010).

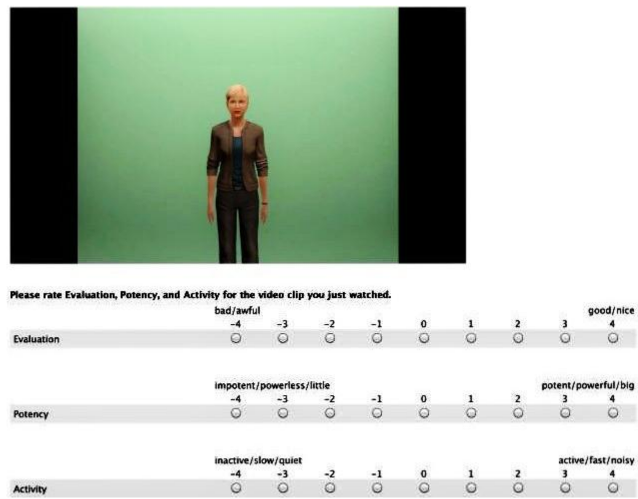

Figure 2. Example Screenshot of the survey showing 'Rachel' the virtual human and the three ratings scales.

We considered survey result data (Malhotra et al. 2014) for total of 90 videos from 27 respondents. The missing entries for each question (E, P, or A) were imputed with the average values across each question. We analyze the consensus of participants across the 6 functional prompts within the 5 emotional styles, first considering consensus across all the data, followed by different subsets.

All Data based Analysis: The Eigenvalue ratios for E were **8.518**, that for P was 1.523 and that for A was 1.914, indicating that the respondents agreed on E dimension.

Gender based Analysis: We then performed analysis based on gender. For males, the Eigenvalue ratios for E, P, A were **6.438**, 1.726, and 1.76 respectively, and for females, the ratios for E, P, A were **8.162**, 1.2, 1.7 respectively, showing higher consensus for E dimension.

Canadian based Analysis: We performed analysis based on responses from Canadian IP addresses only, as the majority (18/27) of the respondents were Canadians. The ratios for E, P, A were **7.148**, 1.522 and 1.61 respectively, indicating that Canadians had agreement on E dimension.

Expected EPA category based Analysis: We analyzed Eigen ratios for each type of expected EPA and the findings were as follows:

+ + +: Eigenvalue ratios for E, P and A were **2.23**, 1.73 and **2.045** respectively, showing more agreement on E and A dimension.

+ - -: Eigenvalue ratios for E, P and A were **2.475**, **2.315** and 1.064, showing more agreement on E and P dimension.

- + +: Eigenvalue ratios for E, P and A were **3.018**, 1.912 and 1.197, showing more agreement on E dimension.

- - - : Eigenvalue ratios for E, P and A as 1.899, 1.713 and **2.003** respectively, showing more agreement on A dimension.

+ - +: Eigenvalue ratios for E, P and A as 1.764, **3.519** and 1.283 respectively, showing more agreement on P dimension.

We see from the above analysis that the first two categories showed consensus in more than one dimension as opposed to all data based, gender based and Canadian based analysis which showed consensus in only E dimension. We computed Pearson's r-values as well, but the results were less conclusive (Malhotra et al. 2014).

4.2 Discussion

The survey analysis showed a consensus in one dimension (E), but not so much in the other two (P and A). This replicates the results originally presented by Osgood (Osgood, George, and Percy 1957; Osgood, May, and Miron 1975) and replicated in many subsequent and cross-cultural studies (Heise 2010): the primary factor that accounts for over half the variance observed in cultural consensus studies is the evaluative one (E), with potency and activity accounting for roughly half as much variance again. In our case, we also have non-verbal behaviors, which leads to further lack of consensus. Further, respondent's comments indicated that they may have been somewhat unsure about how to rate the activity dimension, with some respondents believing it had to do only with the level of motion exhibited by the virtual character. The inherent nature of Potency and Activity dimensions can be one of the causes of ambiguity and hence lack of consensus. For instance, an action displaying more active person can be interpreted as also being powerful.

The results obtained from the survey for each functional step can be incorporated in the enhanced COACH system's prompt selector in different ways. The most straightforward method is to assign an EPA vector for each of these video prompts using the average value from the survey result data and choosing the closest emotional prompt video by calculating the minimum Euclidean distance between the desired emotional EPA vectors (as computed with ACT or BayesAct based on an estimate of the affective identity of the person using the system, e.g. "patient") and the labeled EPA vectors for a specific functional prompt. This method was used in building the system described in (Lin et al. 2014).

However, lack of consensus can be leveraged by the BayesAct engine, as it is a probabilistic model and can evaluate each prompt decision theoretically, using the information about the lack of consensus amongst the respondents. To demonstrate this, let us take a simple example. Consider we have two prompts, P1 and P2. Suppose in our survey that everyone agreed that P1 had an E-value of 0.0, but 50% of the people said P2 had E-value of 2.0 and 50% said P2 had E = -2.0. This shows a lack of consensus amongst our survey population, but also shows that P2 is evaluated differently by different members of the population. For example, we might imagine P2 uses a certain hand gesture that is evaluated very negatively by about half the population due to a cultural difference, whereas P1 uses no hand gesture at all, so is evaluated as quite neutral by everyone. Now, suppose that BayesAct now calls for a prompt with an E-value of 2.0 (very positive), because it has figured out that this is the deflection minimizing prompt. It can pick P1, which it is certain will be evaluated at E=0.0, causing deflection of 2.0. Or, it can pick P2, which will cause zero deflection 50% of the time, and deflection of 4.0 50% of the time. Decision theoretically, the second choice may be a better one. For example, it may be possible to recover from the larger deflection of 4.0 quite easily, and so it is worth the risk. This way BayesAct can model the development of a consensus and a particular user will develop a dyadic consensus about the meaning of certain prompts.

We also compared the mean EPA values measured in the survey with the EPA values that we chose originally for the videos (our expected EPA ratings for the videos). We compared these only as positive/negative agreement for all data and its subsets i.e. Male, Female and Canadian.

EPA+++ and EPA-++ had same expected and actual signs for E, P and A. Complete tables for all the cases can be found in (Malhotra et al. 2014). In the case of expected EPA + - -, the agreement is in E dimension and discrepancies are mostly in dimensions of P and A, i.e. according to respondents the virtual assistant was powerful and active, whereas she was expected to be powerless and inactive. In the case of expected EPA - - -, agreement is mostly in dimensions of E, P and discrepancies are mostly in dimension A, i.e. according to respondents the virtual assistant was active, whereas she was expected to be inactive. In the case of expected EPA + - +, the agreement is mostly in dimensions of E, A and discrepancies are mostly in dimension P, i.e. according to respondents the virtual assistant was powerful, whereas she was expected to be powerless. Overall, there seems to be agreement in E and A dimension from expected EPA sign's perspective.

The free-form comments requested at the end of the survey gave us valuable indications and feedback for future work. The most common comments were focused on the quality of the speech synthesis, which was considered to be "robotic", "not realistic", "too fast" and "not engaging", and the distance that the avatar seemed to be away, such that facial expressions were hard to see. The speech and gestures, and the distance from the camera to the avatar were known limitations of the toolkit, which is a development in progress by the USC group. From a content/delivery perspective of the prompts, some would like less creepiness. We could add music and see its effect.

Other comments indicated that the survey was too long, and that many questions may have been skipped because of this. To remediate, we considered respondents who answered more than 90% of questions. The software we used did not allow for a progress bar to be displayed. We hope to use this feedback given by the respondents to guide our implementation and design choices in future.

5. Conclusion

The user study conducted in this paper has paved a way to improve an assistive handwashing system that helps persons with cognitive disabilities. It will transform the system from being just a hand washing assistant 'system' to being a 'virtual human' assistant having emotions, with the ability to interact through the prompt videos with the users like a human would do. The planned affective hand washing system will use the survey results for the prompt videos and will improve human-computer interaction. Future improvements to the system could be to have the system interact with the user dynamically with prompts, behaviors and facial expressions instead of static videos. We shall explore other virtual agent toolkits specifically tailored to the care setting (Sarel van Vuuren and Leora R. Cherney, 2014). Further survey work will need to be done with the target population (elders with cognitive disabilities) once a refined set of prompts are developed based on the study we present here. We expect significant differences will be found in the responses of participants, mainly due to the wide range and quickly changing affective identities that are held by persons with dementia (Lively and Smith 2011; Orona 1990; Rose Addis and Tippett 2004). Furthermore, the lack of consensus in the P and A dimensions that was found will be a significant issue that needs to be addressed. As we have described, the BayesAct engine can potentially handle this in a near-optimal way. On the other hand, our survey may have included people from many cultures with different interpretations of the prompts (especially the non-verbal components). When designing such a system for elderly persons in a specific culture, we may find more consensus.

References

Borgatti, S. P., & Halgin, D. S. Consensus analysis. A companion to cognitive anthropology, 171-190, 2010.

Fontaine, J.R.J., Scherer, K.R., Roesch, E.B., Ellsworth, P.C.: The world of emotions is not two-dimensional. Psychological Science 18, 1050 – 1057 (2007)

Heise, D. R., Surveying Cultures: Discovering Shared Conceptions & Sentiments, Wiley 2010

Heise, D.R., Expressive Order: Confirming Sentiments in Social Actions. Springer, 2007.

Hoey, J., Schoder, T.: Bayesian Affect Control Theory of Self. AAAI 2015

Hoey, J., Schroder, T., and Alhothali, A.. "Bayesian affect control theory." In Affective Computing and Intelligent Interaction (ACII), 2013 Humaine Association Conference on, pp. 166-172. IEEE, 2013.

Lin, L., Czarnuch, S., Malhotra, A., Yu, L., Schröder, T., Hoey, J. Affectively Aligned Cognitive Assistance using Bayesian Affect Control Theory. In Proc. IWAAL, LNCS 8868, pp. 279–287, 2014.

Lively, K.J., Smith, C.L.: Identity and illness. In: Handbook of the Sociology of Health, Illness, and Healing, pp. 505–525. Springer New York (2011)

MacKinnnon, N.J., Robinson, D.T.: 25 years of research in affect control theory. Advances in Group Processing 31 (2014)

Malhotra, A., Yu, C., Schröder, T., Hoey, J.: An exploratory study into the use of an emotionally aware cognitive assistant. TR# CS-2014-15, U. of Waterloo (2014)

Mihailidis, A., Boger, J., Candido, M., and Hoey, J. The COACH system to assist older adults with dementia through handwashing: an efficacy study. BMC Geriatrics, 8 (28), 2008

Orona, C.J.: Temporality and identity loss due to Alzheimer's disease. Social Science & Medicine 30(11), 1247 – 1256 (1990)

Osgood, Charles E., W.H. May and M.S. Miron. "Cross-Cultural Universals of Affective Meaning. University of Illinois Press, 1975.

Osgood, Charles E., George E. Suci, and Percy H. Tannenbaum. "The measurement of meaning", University of Illinois Press, 1957.

Romney, A. Kimball, Susan C. Weller, and William H. Batchelder. "Culture as consensus: A theory of culture and informant accuracy." American Anthropologist 88.2 (1986).

Rose Addis, D., Tippett, L.: Memory of myself: Autobiographical memory and identity in Alzheimer's disease. Memory 12(1), 56–74 (2004)

Scholl, W.: The socio-emotional basis of human interaction and communication: How we construct our social world. Social Science Information 52, 3 – 33 (2013)

Schröder, T., Rogers, K. B., Ike, S., Mell, J. N., & Scholl, W.. Affective meanings of stereotyped social groups in cross-cultural comparison. Group Processes & Intergroup Relations, 16(6), 717-733, 2013.

Van Vuuren, S. & Cherney, L. R. (2014). A Virtual Therapist for Speech and Language Therapy. In T. Bickmore, S. Marsella, & C. Sidner (Eds.): Intelligent Virtual Agents 14th International Conference, IVA 2014, Lecture Notes in Artificial Intelligence 8637, pp. 438–448. Springer.

Wilson, R., et al.: Quantitative analysis of formal caregivers' use of communication strategies while assisting individuals with moderate and severe alzheimer's disease during oral care. J. of Communication Disorders 46(3), 249–263 (2013)