

# Evaluation of Approaching-Strategies of Temporarily Required Virtual Assistants in Immersive Environments

Andrea Bönsch<sup>1,2,\*</sup>

Tom Vierjahn<sup>1,2,\*</sup>

Torsten W. Kuhlen<sup>1,2,\*</sup>

<sup>1</sup>Visual Computing Institute, RWTH Aachen University, Germany

<sup>2</sup>JARA-HPC, Aachen, Germany



Figure 1: Study overview: (lt) Positions of all items and the starting point per run plotted onto the floor plan of the virtual house. (mid) Embedded assistant in a speech-based conversation with an immersed subject. (rt) An experimenter controlling the assistant in a Wizard-of-Oz paradigm.

## ABSTRACT

Embodied, virtual agents provide users assistance in agent-based support systems. To this end, two closely linked factors have to be considered for the agents' behavioral design: their *presence time* (*PT*), i.e., the time in which the agents are visible, and the *approaching time* (*AT*), i.e., the time span between the user's calling for an agent and the agent's actual availability.

This work focuses on human-like assistants that are embedded in immersive scenes but that are required only temporarily. To the best of our knowledge, guidelines for a suitable trade-off between *PT* and *AT* of these assistants do not yet exist. We address this gap by presenting the results of a controlled within-subjects study in a CAVE. While keeping a low *PT* so that the agent is not perceived as annoying, three strategies affecting the *AT*, namely *fading*, *walking*, and *running*, are evaluated by 40 subjects. The results indicate no clear preference for either behavior. Instead, the necessity of a better trade-off between a low *AT* and an agent's realistic behavior is demonstrated.

**Index Terms:** H.5.2 [User Interfaces]: Evaluation/Methodology—User-Centered Design; I.3.7 [Computer Graphics]: Three-Dimensional Graphics and Realism—Virtual Reality;

## 1 INTRODUCTION

In a wide range of applications, virtual assistants are increasingly common for providing various, situation-dependent and user-centered support as, e.g., interlocutors [6], guides [7] or coaches [4]. Their representation can be either bodiless, e.g., voices from the off, or embodied, i.e., abstract or human-like characters. This work deals with representing assistants as computer-controlled, human-like, virtual agents (VAs). However, designing a believable behavior of these embodied agents is challenging. Besides mirroring social elements of human behavior, e.g., meeting personal space requirements and avoiding collisions with the user and the scene [3], adequate strategies for an assistant to approach users and departing from them have to be modeled. Three potential strategies are in focus of this work.

\*e-mail: {boensch|vierjahn|kuhlen}@vr.rwth-aachen.de

Assistants are required by the user with different frequencies. In most agent-based support systems, the assistants are key components and thus need to be omnipresent. In contrast, certain scenarios exist, where assistants are required only temporarily: they can be called and sent away at the user's discretion.

*Omnipresent* assistants are permanently visible and in almost constant contact with the user. This permanent presence facilitates a quick and unhindered assistance.

This work, however, focuses on *temporarily required* assistants. For these, omnipresent behavior may come with a drawback: as the assistants constantly focus the users, particularly whenever they are not required, they may be quickly perceived as distracting and annoying. Thus, reducing the agent's *PT*, i.e., presenting the assistant only when needed, is desirable. Consequently, the question arises how assistants may return when being required. Each applicable strategy leads to a different *AT*.

Considering temporarily required assistive interfaces like pie menus [5], a common approaching-strategy after triggering a signal of necessity is instantly *fading in*. By this, users can immediately access all available support functions. Furthermore, if the menus are no longer needed, they fade out as a departure-strategy. The resulting low *AT* and *PT* are beneficial and thus suggest themselves as the method of choice for temporarily required assistants.

However, when working with human-like assistants instead of abstract menus, we require realistic behaviors. Due to this, simple appearing (approaching) and disappearing (departure) by immediate *fading* may be disliked. Thus, more realistic approaches may be beneficial, e.g., *walking* or *running*. However, both increase *AT* and *PT*. By conducting a within-subjects study, we investigate which of those three strategies is preferred by users for approaching and departure.

This paper is structured as follows: The design and setup of the user study are described in Section 2 while the results are presented in Section 3 and discussed in Section 4. A conclusion is given in Section 5.

## 2 USER STUDY

We investigate the trade-off between *AT*, *PT* and a VA's pleasant and believable behavior for approaching and departure. To gather the relevant insight, we conducted a user study in a CAVE. We designed two tasks, namely *go-to* and *search*. In these tasks a user has to gather items that are distributed over a scene. Both tasks end with

a user-triggered conversation with the assistant. Three strategies for approaching are tested: immediate fading ( $a_{fade}$ ), walking ( $a_{walk}$ ), and running ( $a_{run}$ ). The departure always mirrors the approaching.

We expect the following hypotheses to be confirmed:

- H1**  $a_{fade}$  is perceived as unrealistic.  
After users trigger a signal of necessity,  $a_{fade}$  ensures the instant appearance of the assistant. We expect users to consider this too unbelievable and unsettling for human-like assistants.
- H2** For the approaching,  $a_{walk}$  and  $a_{run}$  are preferred over  $a_{fade}$ .  
This hypothesis is a direct consequence of H1: although  $a_{fade}$  provides the lowest AT and thus the quickest support, we expect users to tolerate a longer AT in favor of a more human-like VA.
- H3**  $a_{walk}$  is preferred over  $a_{run}$ .  
Our scenario is an indoor-scene, where people are typically walking instead of running around.
- H4**  $a_{fade}$  is preferred as the departure-strategy.  
After the support has ended, we expect the users to focus back on their tasks, ignoring the environment. Therefore, a realistic departure-behavior of the VA is not required anymore. Furthermore, the low PT of fading is beneficial to prevent the assistant from hampering the user.

## 2.1 Apparatus

We used a five-sided CAVE (four walls and floor) with the dimensions  $5.25m \times 5.25m \times 3.30m$  ( $w \times d \times h$ ) providing a  $360^\circ$  horizontal field of regard. The subjects wore active stereo glasses, tracked at 60 Hz. An ART Flystick 2 was provided for navigation and interaction. A loudspeaker array in the CAVE's ceiling was used to play speech-based system commands as well as the assistant's verbal parts. Additionally, a microphone array and two security cameras enabled the experimenter to observe the fully immersed subjects. Provided with a cloned view through the subjects' tracked glasses, the experimenter was thus able to control the assistant in a Wizard-of-Oz paradigm, shown on the right of Figure 1.

## 2.2 Virtual Environment and Task

As shown in the left of Figure 1, our user study scenario is an elaborate, single-story house with five fully-furnished rooms and many items. Subjects had to interact with four of them: a shopping list, a newspaper, a photo, and a note of a medical appointment.

A female character, animated by means of SmartBody [8], is introduced as the computer-controlled assistant. She serves as interlocutor answering the subjects' questions regarding the aforementioned items (see middle of Figure 1). Technical limitations prevented us from using SmartBody's text-to-speech feature. Thus, we pre-generated the required sound files for the conversations using an online tool [1] (settings: US English, voice "Alice", medium speed). By this, the VA's voice, volume and pronunciation did not vary during the user study. For each message, several sound files with different wordings were prepared, facilitating diversified conversations. By using a Wizard-of-Oz paradigm the experimenter selected a situational sound file, resulting in believable conversations, however, lacking lip synchronization.

Subjects had to fulfill two tasks inside the immersive virtual environment (IVE), both ending with a short conversation with the assistant. In the first task, subjects performed a goal-oriented navigation ( $s_{go-to}$ ) to either the shopping list or the newspaper. The positions of both items were announced before the user study. In the second task, subjects performed an explorative navigation ( $s_{search}$ ) to either the photo or the note of the medical appointment. The positions of both items were varied per run to ensure that actual searching was necessary. All positions of the four items and the subject's starting point for each run are plotted in the floor plan on the left of Figure 1. To instruct the subjects which item had to be gathered in the next run, a speech-based approach with a voice from the off was used. The required sound files were generated by the same online tool as before [1] (settings: US English, medium speed). This time we chose voice "John"

to make the instructions clearly distinguishable from the conversations with the female assistant. In order to prevent an incorrect task execution, only the currently announced item was available per run.

After reaching an item, subjects confirmed its detection by ray-casting-based point-and-click approach using the Flystick. The click automatically triggered the appearance of the VA, who was not present during the initial phase of each task. We chose item-confirmation instead of a commercial implementation like calling the agent's name as trigger, to keep the required interaction of our subjects minimal. The assistant appeared using one of the three approaching-strategies, and a short conversation, in which the assistant engaged the subject in eye contact, started. After the assistant had asked an initial question (e.g., "May I help you?"), subjects had to ask a question regarding the corresponding task-item, e.g., where they took the photo. The assistant answered it. After the subjects had thanked for the support, the assistant disappeared, mirroring her previous approaching-strategy in the departure.

## 2.3 Experimental Design and Data Collection

We chose a within-subjects design with two independent variables: (a) the subjects' task ( $s_{go-to}$ ,  $s_{search}$ ) and (b) the VA's behavior ( $a_{fade}$ ,  $a_{walk}$ ,  $a_{run}$ ). At the end of the approaching-strategy, the assistant was positioned within a reasonable distance to the user for the face-to-face conversation. For the walk and run behaviors, the respective start (approaching) and end (departure) points of the assistant's trajectories were located in the neighboring rooms. By this, her sudden appearance and disappearance was not seen by the subjects. Furthermore, the trajectories for approaching and departure varied in order to achieve a more human-like motion. For each behavior, both tasks were tested with both items, thus resulting in gathering all four items per behavior.

We used several data sets to evaluate our hypotheses. Per behavior, questions dealing with the perceived realism of the behavior and the level of satisfaction regarding the assistant's approaching time had to be answered. Additionally, subjects rated their perceived social presence of the assistant by means of the Social Presence Survey (SPS) [2]. After experiencing all three strategies, each behavior had to be rated separately according to certain criteria. In addition, preference on realism versus quick support and the preferred strategy for approaching and departure had to be given. Complementing free-text fields allowed to gain more insight into the subject's experiences. Questions regarding lip synchronization, preferences on variation of speech and trajectories or embodied vs. bodiless speech had to be answered. Besides, the subjects' rated their level of presence by means of the SUS presence questionnaire [9]. Finally, as we expected different preferences regarding the assistant's behavior depending on the time spent for the initial task, we measured the execution times per task.

## 2.4 Procedure

Prior to the study, the subjects were informed about the general procedure. The four items and the positions of the  $s_{go-to}$  items were introduced. Furthermore, the questions subjects had to ask their assistant about the items were presented and had to be memorized. After giving their informed consent, the subjects filled out a demographic questionnaire and entered the CAVE for a familiarization phase. They were asked to explore the house without a time limit. Neither the assistant nor the four items were present in the scene during this familiarization.

After finishing the familiarization phase, the user study began. The execution was divided into three blocks, one per behavior. In each block, subjects had to gather all four items once, resulting in two  $s_{go-to}$  and two  $s_{search}$  task-runs. The three behavior blocks were tested in a randomized order and also the order of  $s_{go-to}$  and  $s_{search}$  in a block was randomized. Each block was followed by a set of questions, which had to be answered by using the Flystick in the IVE.

After leaving the CAVE, subjects filled out a final questionnaire. In total, the user study took about 50 minutes per subject, of which about 24 minutes were spent fully immersed.

Table 1: Results of subjects' ratings per assistant's behaviors regarding three questions on realism and approaching time.

	answer frequencies	M	SD	significances
Realistic behavior	$a_{fade}$	2.93	1.44	}** }+
	$a_{walk}$	4.35	1.53	
	$a_{run}$	3.63	1.58	
More rapid availability in $s_{go-to}$	$a_{fade}$	2.25	1.34	}** }** }*
	$a_{walk}$	4.58	1.71	
	$a_{run}$	3.68	1.58	
More rapid availability in $s_{search}$	$a_{fade}$	2.40	1.37	}** }** }+
	$a_{walk}$	4.45	1.66	
	$a_{run}$	3.70	1.59	

strongly disagree 1 2 3 4 5 6 7 strongly agree

\*\* significant at .001 level, \* significant at .05 level, + non-significant trend at .1 level

## 2.5 Subjects

Forty-four volunteers from the computer science department participated in our study. Due to technical problems regarding the visibility of items or the VA, we discarded data sets of four subjects. Thus, we have a population of forty subjects (32 ♂, 8 ♀, ages  $M=25.5$ ,  $SD=4.49$ ). All of them were naïve to the purpose of the study. All had normal or corrected-to-normal vision, normal hearing, normal motor skills and fluent or at least basic English skills. Twenty-two subjects stated that they had used a CAVE never or only once before, e.g., in a short campus demo. Eighteen subjects worked professionally in a field related to Virtual Reality. Seven subjects had prior experiences with VAs embedded as assistants.

## 3 ANALYSIS AND RESULTS

For evaluation, we used the proposed 7-point Likert scales for the standardized questionnaires (SPS: -3=strongly disagree to 3=strongly agree, SUS: 1 to 7). Our own complementing questions had either a 7-point scale (1 to 7) or several response options (e.g., “yes vs. no” or the three behaviors), from which exactly one had to be chosen.

For all tests, a significance level of 0.05 was used. Furthermore, we analyzed the 7-point scale results as follows: A one-way ANOVA was used, followed by Tukey's HSD post-hoc tests to analyze significant effects. However, when Levene's test indicated that the assumption of homogeneity of variances was violated, Welch's ANOVA was used. When appropriate, Games-Howell post-hoc tests were then used to analyze significant effects.

Subjects were asked to rate their experience per behavior according to three statements on a 7-point Likert scale from 1=strongly disagree to 7=strongly agree, summarized in Table 1. For S1 (*I perceive my assistant's behavior as realistic.*) subjects perceived  $a_{fade}$  as highly significantly ( $F_{2,117}=8.83$ ,  $p<.001$ ) less realistic than  $a_{walk}$ . Besides, a non-significant trend favoring  $a_{walk}$  over  $a_{run}$  with  $p=.087$  was revealed. For S2 (*I expect a more rapid availability of my assistant in the go-to task.*), subjects expected the VA to appear highly significantly ( $F_{2,117}=23.0$ ,  $p<.001$ ) faster in  $a_{walk}$  and  $a_{run}$  compared to  $a_{fade}$ . Furthermore, they expected the VA to appear significantly faster ( $p=.043$ ) in  $a_{walk}$  compared to  $a_{run}$ . Finally, for S3 (*I expect a more rapid availability of my assistant in the search task.*), subjects also expected the VA to appear highly significantly ( $F_{2,117}=18.0$ ,  $p<.001$ ) faster in  $a_{walk}$  and  $a_{run}$  compared to  $a_{fade}$ . In addition, there is a non-significant trend ( $p=.081$ ), favoring a quicker appearance in  $a_{walk}$  compared to  $a_{run}$ .

We computed the SPS score [2] as sum of the five individual SPS ratings per VA's behavior. In order to have a consistent terminology throughout the study, we adapted the items by replacing “person” by “assistant” if applicable. There is no significant difference between  $a_{fade}$  ( $M=5.70$ ,  $SD=4.30$ ),  $a_{walk}$  ( $M=5.18$ ,  $SD=4.94$ ) and  $a_{run}$  ( $M=6.33$ ,  $SD=5.08$ ), with  $M$  denoting the mean and  $SD$  denoting the standard deviation.

Table 2: Results of subjects' ratings per VA's behaviors regarding five opposing statements known from questionnaires for software tools.

	answer frequencies	M	SD	significances
unlikeable	$a_{fade}$	4.43	1.55	}*
	$a_{walk}$	5.08	1.27	
	$a_{run}$	5.13	1.17	
inefficient	$a_{fade}$	5.38	1.39	}**
	$a_{walk}$	3.98	1.80	
	$a_{run}$	4.55	1.84	
obstructive	$a_{fade}$	5.03	1.53	}
	$a_{walk}$	5.03	1.42	
	$a_{run}$	5.03	1.33	
unreliable	$a_{fade}$	5.90	.087	}
	$a_{walk}$	5.50	.96	
	$a_{run}$	5.50	1.20	
human-like	$a_{fade}$	5.03	1.58	}*
	$a_{walk}$	3.68	1.66	
	$a_{run}$	4.28	1.88	

1 2 3 4 5 6 7 abstention

\*\* significant at .001 level, \* significant at .05 level

Table 3: Results of subjects' preference rating on assistant's *realistic behavior vs. quick support* by the assistant.

	answer frequencies	M	SD	signif.
realism of behavior	1 2 3 4 5 6 7	4.03	2.19	

legend: 1 2 3 4 5 6 7

Table 4: Subjects' preferences for approaching and departure.

		behavior			significances	abstention
		$a_{fade}$	$a_{walk}$	$a_{run}$		
approaching	$s_{go-to}$	14	14	12		
	$s_{search}$	14	14	12		
departure	$s_{go-to}$	14	20	5	*	1
	$s_{search}$	13	21	5	*	1

\* significant at .05 level

After experiencing all three behaviors, we asked the subjects for ratings according to pairs of opposing statements known from questionnaires for software tools. Again, a 7-point scale was used. The results are summarized in Table 2. Subjects rated  $a_{run}$  significantly ( $F_{2,77.1}=3.19$ ,  $p=.045$ ) more pleasing than  $a_{fade}$ . They rated  $a_{fade}$  highly significantly ( $F_{2,117}=6.9$ ,  $p<.001$ ) more efficient than  $a_{walk}$ . Finally, they rated  $a_{walk}$  significantly more human-like than  $a_{fade}$  ( $F_{2,117}=6.3$ ,  $p=.002$ ). For the remaining pairs (*obstructive vs. supportive* and *unreliable vs. reliable*), we found no significant differences.

Additionally, we asked the subjects to do a trade-off on a 7-point scale between *the realism of the assistant's behavior* (1) and *quick support* (7) (see Table 3). A One-Sample t-test revealed no significance.

Examining potentially different preferences between approach and departure for  $s_{go-to}$  and  $s_{search}$  was done by means of explicit queries. The results are presented in Table 4. A One-Sample Chi-Square test per row revealed that the frequencies of the preferred behaviors differed significantly across the departure tasks ( $p_{go-to}=.012$ ,  $p_{search}=.007$ ), favoring  $a_{walk}$ .

The lack of the assistant's lip synchronization was not recognized by seven subjects, three didn't give a statement. Judging from some written comments, this missing lip synchronization during the female voice playback had a negative influence on the assistant's realism.

The variation in the conversation was recognized by 28 subjects. 26 of them reported that they liked it. Furthermore, 28 subjects liked the varying trajectories in  $a_{walk}$  and 26 in  $a_{run}$ . One subject abstained from answering both trajectory-related questions. The remaining subjects disliked the variation.

We asked whether the subjects preferred a VA or a voice from the off for speech-based conversations. 35 subjects voted for the VA, 5

for the voice from the off. A One-Sample Binomial Test revealed a highly significant difference between both ratings ( $p < .001$ ).

The subjects rated the feeling of being present in the IVE as reasonably high. An average SUS score of  $M=4.55$  ( $SD=.705$ ) was reported.

Finally, we evaluated the execution times per task:  $M_{go-to}=13.9$  s ( $SD_{go-to}=8.01$  s),  $M_{search}=54.7$  s ( $SD_{search}=44.1$  s). Thus, the mean execution time of  $s_{search}$  is four times larger than the one of  $s_{go-to}$ .

## 4 DISCUSSION

According to the task description (Section 2.2), the interaction between the VA and the subjects stayed the same throughout the study. Thus, differences in the subjects' preferences can be traced back to the influence of approaching- and departure-behavior.

Although the female assistant is computer-controlled and thus a technical user interface, her visual appearance is human-like. According to the subjects' ratings, the behaviors  $a_{run}$  and especially  $a_{walk}$  further support her human-likeness. This finding is backed up by the subjects' clear rating of  $a_{fade}$  being technical, i.e., unrealistic. Thus, **H1** is confirmed.

With  $a_{fade}$  being rated as unrealistic, we expected it to be an unfavorable choice for an assistant's approaching-behavior. Instead, the more realistic  $a_{walk}$  or  $a_{run}$  might be more beneficial. However, none of the three behaviors was preferred by our subjects for the approaching, contradicting **H2**. Nevertheless, we received different results when considering the departure. Here, subjects clearly preferred  $a_{walk}$  over  $a_{run}$  and  $a_{fade}$ , contradicting **H4**. Interestingly, the different mean execution times of the tasks  $s_{go-to}$  and  $s_{search}$  had no influence on the subjects preferences, neither for approaching nor for departure. This indicates that even if users might already be slightly annoyed after a time-consuming search – as pointed out by some subjects – the preference regarding the assistant's behavior does not change.

Taking all the results given in Section 3 into account, **H3** can only be confirmed for the departure. For the approaching, we received mixed feedback. Statements like "I feel strange with running at home." or running "makes me nervous" and I feel "stressed", were given as reasons to refuse  $a_{run}$ . In contrast, it was also reported that  $a_{run}$  gives "the user the feeling of being important", which we consider beneficial for an assistive interface. Thus, more research needs to be done on users' preferences regarding the realistic strategies.

In fact, the user study could not reveal a clear preference towards either a realistic or a non-realistic approaching-behavior. By this, it is highly likely that there are no universal design guidelines which are accepted by all users in all circumstances. However, some interesting information can still be derived, especially when taking the results from the free-text fields on the questionnaires into account.

An assistant's primary goal is to support users. Traits like being supportive or being reliable were rated similarly high for all three behaviors, without any significant difference. Thus, we conclude that the approaching-strategy does not effect the VA's utility.

If users have to wait too long for the assistant's support, they "can also do the task on their own", as stated by one of our subjects. Thus, assistants need to support users efficiently. In our user study, the subjects perceived  $a_{fade}$  as the most efficient technique, to such an extent that they refused the need for a more rapid availability of the assistant, e.g., by avoiding the initial question. In contrast, subjects expected a quicker availability for the running and the walking assistant. Thus, the AT is a crucial factor for the perceived efficiency and should be kept as low as possible. This is supported by further statements given by the subjects, e.g., "Assistants shouldn't include disadvantages of reality, efficiency is more important." or "Assistants are still computer interfaces and I don't want to spend more time than necessary waiting for a computer task."

However, other subjects voted that a realistic behavior of the assistant should be prioritized: "Realism over technical zero-delay response." or arguments like "If there is too little focus on realism, the assistant has no point" and "Realism is the main priority in

applications with focus on VR" are given. In accordance with these statements, also the rating whether the realism of a human-like assistant should have higher priority or the quick support provided by this assistant, indicates no clear preference. Instead, subjects asked for a better trade-off between both crucial factors, supporting our primary expectation and motivation for the user study. However, evidently more research has to be done in this direction.

Finally, also the attractiveness of the assistant turned out to be a key factor for the users' acceptance of the VA. (1) Our results indicate that lip synchronization plays an important role for a convincing VA's performance. (2) Variations in speech as well as in walking trajectories are recommended. (3) Interestingly, neither realistic nor unrealistic behavior seem to influence the subjects' perception of the social presence of the VA. (4) However, the behaviors have an influence on the assistant's characteristic of being pleasing. Here, the assistant showing a realistic behavior was rated better. (5) The subjects preferred having a visual VA representation during speech-based conversations, independent of approaching and departure. This again suggests that research for improving the design of virtual assistants needs to be done.

## 5 CONCLUSION

We presented an evaluation of three strategies ( $a_{fade}$ ,  $a_{walk}$  and  $a_{run}$ ) for the approaching and departure of a temporarily required, human-like, virtual assistant. While the subjects perceived  $a_{fade}$  as clearly unrealistic, it was not considered an unfavorable approaching-behavior due to its low AT. Interestingly, we found no clear preference towards a single approaching-behavior. Only for the departure,  $a_{walk}$  was preferred, probably due to the high level of realism.

Thus, we cannot yet give a clear recommendation for the design of assistants who are required only temporarily. Instead, more research on a suitable trade-off between AT, PT and an assistant's realism has to be done. Thereby, characteristics like efficiency, attractiveness, e.g., in form of being pleasing, or being human-like, need to be evaluated.

In future work, we plan to address this open issue. Therefore, we intend to complement our behavior design with the behaviors *omnipresence* or *working self-reliantly in the user's vicinity*, while using an improved task for the user and the assistant to commonly work on.

## ACKNOWLEDGEMENTS

This work was funded by the project house *ICT Foundations of a Digitized Industry, Economy, and Society* at RWTH Aachen University.

## REFERENCES

- [1] From Text To Speech - Free online TTS service. <http://www.fromtexttospeech.com>, last visited 2016-10-18.
- [2] J. N. Bailenson, J. Blascovich, A. Beall, and J. Loomis. Equilibrium Theory Revisited: Mutual Gaze and Personal Space in Virtual Environments. *Presence*, 10(6):583–598, 2001.
- [3] A. Bönsch, B. Weyers, J. Wendt, S. Freitag, and T. W. Kuhlen. Collision Avoidance in the Presence of a Virtual Agent in Small-Scale Virtual Environments. In *IEEE Symp. on 3D User Interfaces*, pages 145–148, 2016.
- [4] I. de Kok, J. Hough, F. Hülsmann, M. Botsch, D. Schlangen, and S. Kopp. A Multimodal System for Real-Time Action Instruction in Motor Skill Learning. In *Proc. of the 2015 ACM on Intern. Conf. on Multimodal Interaction*, pages 355–362, 2015.
- [5] S. Gebhardt, S. Pick, F. Leithold, B. Hentschel, and T. W. Kuhlen. Extended Pie Menus for Immersive Virtual Environments. *IEEE Trans. on Visualization and Computer Graphics*, 19(4):644–651, 2013.
- [6] S. McGlashan. Speech Interfaces to Virtual Reality. In *Proc. of the 2nd Intern. Workshop on Military App. of Synthetic Environments and VR*, 1995.
- [7] A. Roque, D. Jan, M. Core, and D. Traum. Using Virtual Tour Behavior to Build Dialogue Models for Training Review. In *Proc. of the 10th Intern. Conf. on Intelligent Virtual Agents*, pages 100–105, 2011.
- [8] A. Shapiro. *Motion in Games*, volume 7060, chapter Building a Character Animation System, pages 98–109. 2011.
- [9] M. Usoh, E. Catena, S. Arman, and M. Slater. Using Presence Questionnaires in Reality. *Presence*, 9(5):497–503, 2000.