Interpreting Social Commitment in a Simulated Theater

Daniel Veutgen
University Duisburg-Essen
daniel.veutgen@stud.uni-due.de

Marco Massetti
Joy Rossi
Leonardo Veroli
University of Camerino

Athena Ásgeirsdóttir
Gudrun Baldursdóttir
Sigrún Gissurardóttir
Gísli Gudmundsson
Tinna Sigurdardóttir
CADIA, Reykjavik University

Valdimar Laenen
Amsterdam University College

Hannes Högni Vilhjálmsson
CADIA, Reykjavik University

ABSTRACT

Awareness of the surroundings as well as bodily cues is essential for effective social interaction. Nonverbal behavior such as gaze, facial expression, gesture and posture, provide cues for interpreting intent and relational status. These cues are also present as people traverse a physical setting, where locomotion and social expression integrate effortlessly. The goal of this study was to examine how human observers interpreted social behavior of virtual agents constrained by the physical environment of a movie theater with seating rows. An on-line survey compared videos of agents exhibiting only locomotion, and agents that also exhibited two different levels of social commitment. Results showed that people were able to interpret the social behavior correctly, indicating that adding it on top of complex maneuvering is possible. These results continue to build a bridge between the literature on agent navigation and agent social behavior.

CCS CONCEPTS

• **Computing methodologies** → **Intelligent agents**; **Procedural animation**; *Agent / discrete models*;

KEYWORDS

Virtual Agents, Navigation, Social Commitment, Social Awareness

ACM Reference Format:
Daniel Veutgen, Marco Massetti, Joy Rossi, Leonardo Veroli, Athena Ásgeirsdóttir, Gudrun Baldursdóttir, Sigrún Gissurardóttir, Gísli Gudmundsson, Tinna Sigurdardóttir, Valdimar Laenen, and Hannes Högni Vilhjálmsson. 2018. Interpreting Social Commitment in a Simulated Theater. In *IVA '18: International Conference on Intelligent Virtual Agents (IVA '18), November 5–8, 2018, Sydney, NSW, Australia.* ACM, New York, NY, USA, 6 pages. https://doi.org/10.1145/3267851.3267919

IVA '18, November 5–8, 2018, Sydney, NSW, Australia

ACM ISBN 978-1-4503-6013-5/18/11…$15.00
https://doi.org/10.1145/3267851.3267919

1 INTRODUCTION

When humans interact with one another, awareness of the surroundings as well as bodily cues is essential for effective communication. Nonverbal behavior such as gaze, facial expression, gesture and posture, provide cues for interpreting intent and relational status [8]. These cues are also present as people traverse a physical setting, where the interplay between navigation, environmental constraints and social expression weave a layered fabric of behavior, masterfully coordinated, yet without much conscious effort. While maneuvering a space, people can simultaneously express the strength of relationship by the way they carefully manage their inner distances and walking pace.

For interactive virtual agents achieving a similar level of seamless integration of human-environment and human-human behavior is a challenge, calling for careful evaluation of the effectiveness of social cues under environmental constraints. So far, little research has been done to evaluate the degree to which social commitment is readable in a complex social scene, both during navigation and agent-to-agent interaction. Typical implementations of navigation tend to solve path-finding and obstacle avoidance but sometimes lack the depth to address the social dimension of the navigation context. While we have started to develop navigation algorithms that consider social spaces [13], further evaluation is needed.

The primary focus of this work is to evaluate a relatively simple implementation of social commitment in the physical setting of a movie theater, in an attempt to motivate further work into defining a minimal collection of behaviors reusable in both agent navigation and social group interaction. Our goal is to gain a better understanding of whether a human observer can read a spectrum of agent behaviors communicating low commitment, high commitment, or no social awareness at all. An on-line survey with video stimuli and questionnaires was conducted to answer these questions.

2 RELATED WORK

2.1 Social Commitment

People rely on nonverbal behavior to communicate their level of participation in a group activity. People can exhibit a high level of participation, a low level, or no participation at all. When people participate in a social activity their body posture and gaze behavior reflects a degree of attention and engagement in the activity. This level of participation is in fact a declaration of commitment to the activity itself. For example, think of situations in which you are

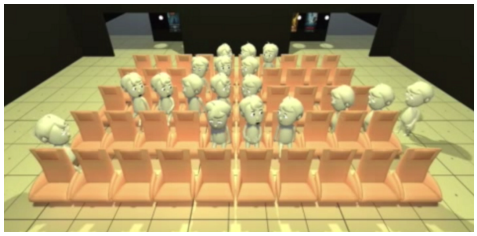

Figure 1: The scenario is a movie theater filling with people.

talking to a group of friends and you are all standing close to each other and pointing your feet and torso towards them. Your phone buzzes after receiving a message. You might check your phone but trying to keep it short because your posture broadcast to everyone the high social commitment you show towards your friends and pressures you to engage back with the group activity as soon as possible. Scheflen [15] divides the human body into four regions which people typically use to communicate a degree of commitment to a social activity. By orienting certain body regions, people can express their stance and interest towards a conversation. Scheflen distinguishes between a low-commitment configuration, which is for example characterized by the commitment of just one body region, and a high-commitment configuration, where more body regions are committed [15]. Scheflen observed a variety of such patterns that characterize commitment, summarized in Table 1.

2.2 Agent Navigation and Awareness

Virtual agents have become quite successful with navigation tasks like path finding and collision avoidance. Therefore, the focus of scientific research has shifted a bit in recent years. The navigation of groups [7] and path finding with other social entities in the environment [13, 14] have became more interesting for research. With this focus, social aspects of the physical environment are being taken into account. For a credible implementation, agents must be equipped with some sense of social awareness [14]. Generally this may be considered the ability to take a different perspective [9, 16]. Considering the perspective of others allows us to predict behaviors and concerns [3], develop empathy and social sensitivity [16] and moreover to perform social skills like solving conflict or negotiate with others [9].

3 SCENARIO

In order to experiment with the interplay between physical constraints, imposed by the environment, and social behavior, we chose a movie theater as a setting (see Fig. 1). We built humanoid agents that enter the theater, either by themselves or in groups of a maximum of three friends. Three settings were implemented for the agents: (1) No Social Awareness; (2) Social Awareness, but with Low Social Commitment; and (3) Social Awareness, with High Social Commitment. These correspond to the three conditions we were interested in analyzing and comparing. The reason for this split is that first we wanted to establish a baseline human-environment behavior, where the agents demonstrated a relatively complex locomotion skill (i.e. finding and getting into their seat). Then on top of this baseline behavior, we "turned on" social awareness, which gave them the ability to sense and respond to other agents socially. The two kinds of social behavior, levels of commitment, were created to see whether subtle social behavior difference was interpretable, on top of the already complex baseline. The implementation of this simulation is described in detail in Massetti et al. [11], but here we summarize how locomotion and social behavior were approached.

3.1 Basic Locomotion

The baseline human-environment interaction, involved implementing proper navigation from the cinema entrance to the seats. This consisted of typical path-finding and obstacle avoidance extended with the ability to walk in closer groups (for agents friends) and the ability to walk down seating rows. To allow an agent to squeeze past seated guests we marked the space in front of every row in special ways using a technique akin to smart zones which informs the agent of the appropriate posture to maintain given certain conditions in that particular location.

The seat the agents chose depended on three factors: (1) Value of the seat (best towards center); (2) Whether they entered the cinema alone or in a group; and (3) The level of commitment and awareness. If the agent is alone, it just chooses the highest valued seat. If they are in a group, they follow a group leader that chooses seats with the highest value for the whole group. The level of commitment is the third factor influencing the choice. In the low commitment and no awareness conditions, they will try to choose a spot with at least one empty seat between them and the next agent, whereas in the high commitment condition it makes no difference whether the next seat is occupied. The interaction and orchestration was implemented by using a combination of inverse kinematics and coordination logic [11].

Once seated, the agents that came with friends were instructed to start a conversation until the movie started. These conversations had to be conducted while sitting in front-facing seats, placing serious restrictions on the typical formation for group conversations.

3.2 Social Behavior

Our agents can express social behavior by means of gaze, facial expressions and hand gestures. The gaze controller allows precise duration specification and synchronization with other behaviors. We referred to some of Argyle's observations on gaze patterns [1] and differentiated the duration of gazes between listeners and talkers in conversation. When the agent itself was moving around, its gaze was directed to points on its path. We coordinated gaze with other gestures to achieve more complex behavior such as communicating low commitment by holding a mobile in the hand and looking down at the phone. Facial expression was stylized and made by combining the movement of the eyes, eyelids, eyebrows and mouth. The degree of expressiveness varied from maximum in the high commitment condition to a neutral expression in the no awareness condition (see Fig. 2).

Even though the mouth is moving, there is no actual speech in the scene. The agents are able to show five different emotional expressions (anger, happiness, sadness, uncertainty and surprise) as well as a wink demonstrating complicity. Moreover, a default face was shown when there was no active functional behavior. Beside the facial expressions, additional features such as distance

Table 1: Characteristics of high and low commitment according to Scheflen & Ashcraft [15] (page 73)

	High Commitment	Low Commitment
1	the commitment of a number of body regions;	committing one body region;
2	the maintenance of a minimal interpersonal distance;	maintaining a maximum distance from the focus of activity;
3	the orienting of body regions at a minimal angle of orientation	orienting body regions at an angle such that they are only partly pointed toward the focus;
4	the uncrossing of arms and legs and the use of stances that define a sharp and excluding channel of space;	crossing arms and legs or otherwise covering regions of the body;
5	an active use of mutual point units[1] and multiple connections and their enactment in an expressive heightened style.	keeping body regions inactive or immobile and thus performing a minimum of behavioral manifestations.

[1] Pointing action

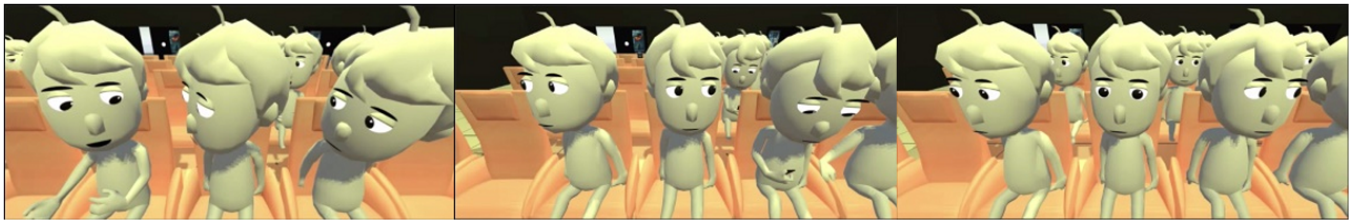

Figure 2: Front view of agents in the high commitment, low commitment and no awareness condition (from left to right).

salutation, showing of endorsement and disapproval, farewell and request for feedback were added [11].

In the high commitment condition the agents are more willing to engage in interaction with others than in the low commitment condition. During the conversation in a group of three highly committed agents, two of them look at the speaker. If the speaker is not in the center position, the listening agent on the opposite position leans in to be able to participate fully. In contrast, in the low-commitment condition agents that enter the cinema alone avoid any conversation. If one member in a low-commitment group of agents starts a conversation, the other two members are listening without any attempt to intervene. Moreover, their facial expression will be mainly negative, and they get easily distracted by their smart phones. In the no awareness condition, no conversation takes place.

4 METHOD

4.1 Procedure

To analyze how people interpreted the behaviors generated by the agents, participants, recruited through mailing lists and social media, were asked to watch three different videos in a web survey. The sequence of the videos was randomized during the survey. Each video took between 49 and 54 seconds and showed the agents entering a movie theater with five rows of seats with ten seats each, looking for a place to sit, followed by a front medium shot with three agents sitting. Fig. 2 displays this last camera angle for all three videos. The difference in the three videos were the display of awareness and commitment by the agents: high commitment[1], low commitment[2] and no awareness[3].

[1]See video at: https://tinyurl.com/high-commitment
[2]See video at: https://tinyurl.com/low-commitment
[3]See video at: https://tinyurl.com/no-awareness

4.2 Measurements

Participants filled in several questionnaires after watching each video. There is no established instrument for assessing the perceived commitment of virtual agents to each other by a human observer. We adapted instruments used by other researchers, focusing on other social constructs, expecting more reliable results than by creating our own instruments from scratch. To ensure the reliability of the adopted questionnaires, Cronbach's α was calculated for the scales. Every scale we used showed a Cronbach's $\alpha > 0.7$, which is the suggested minimum value of reliability when comparing groups [2]. We did not differentiate between entering the cinema and the agent interaction during the questionnaires, because our goal was to create an overall believable scene. This includes the spectator's impression of both behavioral sequences.

The first questionnaire was the Temple Presence Inventory (TPI) [10]. The TPI is used to measure (tele-) presence among 6 dimensions: spatial presence, social presence (actor within medium, passive interpersonal, active interpersonal), engagement, social richness, social realism and perceptual realism. Since some dimensions address VR-experiences, such as spatial presence, social presence (actor within medium) and perceptual realism, they were not included. The questions, that were selected for each dimension are displayed in Table 2. Since a VR version of the experiment is planned, using the TPI helps with a basic comparison.

In the next part of the survey, participants rated the quality of the agent interaction based on a questionnaire that was used to measure the quality of turn taking. It was assessed along the bipolar 5-point Likert scale used by [17]. In total, the turn taking measurement contained 13 adjectives. Bipolar adjectives were for example "Unfriendly vs. Friendly", "Distant vs. Pleasant".

The Perceived Other's Copresence Scale by [12] measures self-reported copresence, perceived others's copresence, telepresence

Table 2: Selected questions from the Temple Presence Inventory

Dimension	Type	Questions
social presence - passive interpersonal	7 pt Likert	During the media experience how well were you able to observe the facial expressions of the people you saw? During the media experience how well were you able to observe the body language of the people you saw?
social presence - active interpersonal	7 pt Likert	How often did you smile in response to someone you saw in the media environment? How often did you want to or did you speak to a person you saw in the media environment?
social presence - engagement (mental immersion)	7 pt Likert	To what extent did you feel mentally immersed in the experience? How involving was the experience? How completely were your senses engaged? To what extent did you experience a sensation of reality? How relaxing or exciting was the experience? How engaging was the story?
social richness	Bipolar Scale	Remote - Immediate; Unemotional - Emotional; Unresponsive âĂŞ Responsive; Dead - Lively; Impersonal - Personal; Insensitive - Sensitive; Unsociable - Sociable
social realism	Agreement with Statement	The events I saw have occurred in the real world. The events I saw could occur in the real world. The way in which the events I saw occurring is a lot like the way they occur in the real world.

and social presence. The telepresence measures the amount of involvement in a VR scene and social presence the social richness of a medium itself. Therefore, only the subscales of self-reported copresence, perceived other's copresence and telepresence were used in this survey. Social richness was already assessed during the TPI. Since the participants were not part of the conversation they observed, the subscales were adapted further, and the questions were reworded from a first-person perspective to a third-person perspective. That is, for example "My interaction partner acted bored by our conversation" became "The interaction partners acted bored during the conversation". Participants rated the statements along a 5-point Likert scale from "Strongly agree" (1) to "Strongly Disagree (5)". Some items of the scale have to be reverse coded for statistics.

The perceived awareness of the agents in the video was assessed with an adapted version of the Mindful Attention Awareness Scale (MAAS) [5]. The original MAAS contains 15 Items which are rated on a 6-point Likert scale from "Almost always" to "Almost never". A higher score mirrors a higher level of dispositional mindfulness. The original MAAS is a first-person perspective questionnaire to assess own awareness. Instead, we were interested if the agents were perceived as aware. Therefore, we adapted MAAS to what we call Mindful Other's Attention Awareness Scale (MOAAS), by changing it to a third-person perspective: Instead of "I rush through activities without being really attentive to them" we asked participants to rate the statement "The persons rush through activities without being really attentive to them". From the 15-item questionnaire, we used 6 items (MAAS 1, 3, 4, 7, 8, 14) for this survey.

Lastly, participants rated the default face of the video agent. They were to rate the similarity of the face to the basic emotions by [6] on a 5-point Likert scale from "agree strongly" to "disagree strongly". Moreover, the participant rated whether the face "appears neutral to me". This was asked to exclude the agent default face as a possible confounding variable, thus detecting a potential non-neutral perception. Emotional stimuli can lead to emotional answers [4]. Therefore, the impression of the stimulus material is important for later interpretation of the results.

Figure 3: Default face rated by participants.

5 RESULTS

In total, 37 participants took part in the approximately 15 to 20-minute survey. Participants came mainly from Europe, such as Italy, France, Denmark and Germany. But there were also participants from the United States and Taiwan. 19 participants were in the 18-24 age group and the others in the 25-34 age group. Among the participants, 18 identified themselves as female and 19 as male. 60.4% played video games less than few times per month (once in a while or never), 2 participants played every day, 9 a few times per week and 4 a few times per month.

Nearly all questionnaires of the survey showed a significant effect, indicating that human observers were able to recognize the different agent behavior in the different video conditions. Only two

Table 3: Mean values and standard deviations of measured dimensions.

	No Awareness		Low Commitment		High Commitment	
	M	SD	M	SD	M	SD
TPI - Social Presence - Passive Interpersonal	9.32	3.46	9.92	3.32	10.16	3.27
TPI - Social Presence - Active Interpersonal[1]	3.11	1.82	3.49	1.94	3.84	2.33
TPI - Social Presence - Engagement [2]	13.73	7.05	16.49	8.15	18.11	8.02
TPI - Social Richness[2]	16.19	6.42	25.62	9.10	30.64	9.68
TPI - Social Realism[1,3]	11.73	5.33	13.68	5.00	14.14	4.72
Interaction quality/ TurnTaking[2]	29.24	6.78	36.19	8.16	43.95	8.10
MOAAS[2]	17.14	6.96	18.59	6.30	24.35	6.16
Self-Reported Copresence[2]	25.51	3.78	20.86	5.00	13.65	3.95
Perceived Other's Copresence[2]	48.14	9.03	43.54	9.00	29.08	8.89
Telepresence[2,3]	9.57	4.38	11.54	5.81	14.21	7.27

[1] $p < .05$

[2] $p < .001$

[3] significant Mauchly's Test of Sphericity

questionnaires showed no significant effect. All mean values and standard deviations are displayed in Table 3.

Four out of the five dimension of the TPI showed a significant effect. Only the repeated measures ANOVA of the passive interpersonal dimension showed no significant effect $F(2,72) = 1.478$, $p = .235$. Active interpersonal showed a significant difference between the video conditions ($F(2,72) = 4.22$, $p = .019$) as well as Engagement ($F(2,72) = 13.04$, $p < .001$), Social Richness ($F(2,72) = 45.81$, $p < .001$) and Social Realism ($F(1.62,58.39) = 5.74$, $p = .009$). Mauchly's Test of Sphericity was significant for Social Realism, indicating that the assumption of sphericity had been violated $\chi^2(2) = 9.29, p = .01$. Therefore, Greenhouse-Geisser is reported. For all significant TPI dimensions, high commitment has the highest mean value, followed by the low commitment condition. The no awareness condition has the lowest mean value. Therefore, we can assume the video with the highly committed agents is displaying the greatest social presence in terms of active interpersonal and engagement as well as social richness and realism.

Sphericity was also violated for the Telepresence dimension of the adapted Perceived Other's Copresence Scale $\chi^2(2) = 12.7, p = .002$). For this reason, Greenhouse-Geisser is reported, which showed a significant effect ($F(1.53,55.2) = 16.15$ $p < .001$) with highest mean value for the high commitment condition and lowest mean value for no awareness. The mean value of the low commitment condition is in between. Since in the other two dimension no violation of sphericity was detected, the regular F value is reported. The repeated measures ANOVA showed a significant effect for both, Self-Reported Copresence with $F(2,72) = 81.31$ and $p < .001$ as well as for Perceived Other's Copresence with $F(2,72) = 49.62$ and $p < .001$. Considering the rating of the scales for both dimensions was from 1 ("Strongly agree") to 5 ("Strongly disagree"), a lower mean value indicates a higher level of approval. No awareness showed the highest mean value, equivalent with lowest consent. The lowest mean value was found for high commitment, followed by low commitment where the mean value was more than 7 points higher for self-reported copresence and more than 14 points higher for perceived other's copresence.

For Turn Taking (interaction quality) and the MOAAS Mauchly's Test of Sphericity was also not significant, and sphericity can be assumed. The repeated measures ANOVA showed a significant effect for Turn Taking ($F(2,72)=60.16$, $p<.001$) as well as MOAAS ($F(2,72) = 21.71$, $p<.001$). Just like the other measurements, mean values supported the strength of high commitment over low commitment, with the second highest mean value, as well as over the no awareness condition.

Post-hoc tests calculated for the significant dimensions of the repeated measures ANOVA supported our assumptions. We had significant differences between all three conditions in Social Richness in the TPI (all $p < .008$), Telepresence (all $p < .013$), Self Reported Copresence (all $p < .001$), and Turn Taking (all $p < .001$). We found significant differences between the no awareness condition and low commitment condition as well as between the no awareness condition and the high commitment condition, but not for low commitment versus high commitment, for TPI Engagement ($p = .008$ and $p > .001$) and Social Realism ($p = .016$ and $p = .038$). For Perceived Other's Copresence and MOAAS, we found significant effects for the no awareness compared to high commitment group as well as for the low commitment versus high commitment group, but not for no awareness versus low commitment. The significant group differences were $p < .001$.

Finally, we were interested in how participants rated the agent's default face that can be seen in Fig. 3. Most participants rated the face either as sad or neutral. In both cases, 9 participants "agreed strongly", and 16 participants "agreed". The mean values were equal ($M = 2.22$), whereas the standard deviation was slightly higher for sad ($SD = 1.03$) compared to neutral ($SD = .98$). A comparison between sad and neutral showed no significant effect with $F(1,36) < .01$ and $p = 1.0$, indicating that participants recognized the face as sad and neutral to the same amount.

6 DISCUSSION

The main question which led us to conduct this survey was whether humans would recognize the difference between agents that were competently navigating an environment and agents that additionally were acting in a social way, either showing low commitment or high commitment. Since nearly all measurements showed a significant effect, we found strong support for the assumption that participants were able to spot differences in the video stimuli. Moreover, the mean values indicate more perceived credibility and sociality for the high commitment condition. The high commitment video was rated as most engaging, socially rich and realistic. Agents in this condition were associated with a higher quality of interaction, dispositional mindfulness, copresence and telepresence.

The only measurement without a significant effect was the passive interpersonal dimension of the TPI. Considering the questions emphasize mainly whether participants are able to observe facial expression and body language, the non-significant result revealed that there was no difference regarding the quality of the medium. Consequently, we can exclude the video material itself as a potential influential factor causing the other significant results.

Participants noticed not just a difference between the no-awareness condition and any social condition. Apparently, they were able to differentiate clearly between the two levels of social behavior. Both commitment conditions were implemented based on Scheflen's observations regarding body language and posture [15]. Applying these behaviors to stylized characters, with larger heads and eyes than normal humans, does not seem to interfere with the expected interpretation.

When evaluating the agent's default face without emotion or talking, participants rated it both sad and neutral. Considering that participants were asked to rate the face along eight emotions (including neutral), the results indicate that the default face is not causing too much confusions regarding its expression. Nevertheless, we didn't expect the face to be perceived as sad. This may have influenced the perception of the videos in general. On the one hand, participants could have associated the sad face with more empathy. In this case, videos where agents showed the default face longer and more often, like in the no-awareness condition, should have been perceived as more social. On the other hand, the sad face might be associated with malaise and depression and the assumption that the agents feel uncomfortable in the conversation, which could have led to less believability in the scene. Normally, people going to the cinema are not sad to be there. This might need further investigation.

7 CONCLUSION

The goal of this study was to examine how human observers interpreted social behavior constrained by the physical environment, specifically in the setting of a movie theater with seating rows. We found strong evidence that the social behavior rose above the complex maneuvering and posturing imposed by the environment, sending its own clear message. Moreover, it appears that participants were able to differentiate between different types of social behavior, indicating that it is possible to modulate the social behavior on top of the baseline movement and sitting configuration. These results continue to build a bridge between the literature on agent navigation and agent social behavior, encouraging their integration.

Some limitations should be mentioned. Online surveys prevent us from knowing how carefully subjects watched the videos and the instruments used limit us to quantitative data. It is also hard to control for English language fluency, which could impact concept interpretation. These limitations will be addressed in future work where we will conduct the experiment again in a controlled lab setting, offering participants to experience the scene first-hand in virtual reality.

ACKNOWLEDGMENTS

The authors would like to thank Claudio Pedica for his important contribution to the social agent simulation and valuable guidance during the project. This work was supported by The Icelandic Research Fund under Grant Number 174444-051.

REFERENCES

[1] Michael Argyle and Roger Ingham. 1972. Gaze, Mutual Gaze, and Proximity. *Semiotica* 6, 1 (1972), 32–49. https://doi.org/10.1515/semi.1972.6.1.32
[2] J. M. Bland and D. G. Altman. 1997. Statiustics Notes: Chronbach's Alpha. *BMJ* 314, 7080 (1997), 572.
[3] Ray Esteban Boyatzis. 2008. Competencies in the 21st century. *Journal of management development* 27, 1 (2008), 5–12.
[4] T Brosch, G Pourtois, and D Sander. 2010. The perception and categorization of emotional stimuli: A review. *Cognition and Emotion* 24, 3 (2010), 377–400. https://doi.org/10.1080/02699930902975754
[5] Kirk Warren Brown and Richard M. Ryan. 2003. The Benefits of Being Present: Mindfulness and Its Role in Psychological Well-Being. , 822–848 pages. https://doi.org/10.1037/0022-3514.84.4.822 arXiv:NIHMS150003
[6] Paul Ekman. 1992. An Argument for Basic Emotions. *Cognition and Emotion* 6, 3-4 (1992), 169–200. https://doi.org/10.1080/02699939208411068 arXiv:a
[7] Tianyu Huang, Mubbasir Kapadia, Norman I. Badler, and Marcelo Kallmann. 2014. Path planning for coherent and persistent groups. In *Proceedings - IEEE International Conference on Robotics and Automation*. 1652–1659. https://doi.org/10.1109/ICRA.2014.6907073
[8] Mark L Knapp and Judith A Hall. 2009. *Nonverbal Communication in Human Interaction*. Vol. 5. 496 pages. http://books.google.com/books?id=j5HIlfRUPm0C
[9] Alison Lobron and Robert Selman. 2007. The Interdependence of Social Awareness and Literacy Instruction. *The Reading Teacher* 60, 6 (2007), 528–537. https://doi.org/10.1598/RT.60.6.3
[10] Matthew Lombard, Theresa B Ditton, and Lisa Weinstein. 2009. Measuring Presence: The Temple Presence Inventory. *Proceeding of Presence 2009 : the 12th International Workshop on Presence* (2009), 1–14.
[11] Marco Massetti, Joy Rossi, and Leonardo Veroli. 2017. *Territoriality and Visible Social Commitment for Virtual Agents*. Master's thesis. University of Camerino and Reykjavik University, Camerino, Italy.
[12] Kristine L. Nowak and Frank Biocca. 2003. The Effect of the Agency and Anthropomorphism on Users' Sense of Telepresence, Copresence, and Social Presence in Virtual Environments. *Presence: Teleoperators and Virtual Environments* 12, 5 (2003), 481–494. https://doi.org/10.1162/105474603322761289
[13] Carmine Oliva and Hannes Högni Vilhjálmsson. 2014. Prediction in social path following. In *Proceedings of the seventh international conference on motion in games*. ACM, New York, NY, USA, 103–108.
[14] Claudio Pedica and Hannes Hogni Vilhjalmsson. 2010. Spontaneous Avatar Behavior for Human Territoriality. *Appl. Artif. Intell.* 24, 6 (July 2010), 575–593. https://doi.org/10.1080/08839514.2010.492165
[15] Albert Emil Scheflen and Norman Ashcraft. 1976. *Human Territories: How We Behave in Space-Time*. Prentice Hall.
[16] K. M. Sheldon. 1996. The Social Awareness Inventory: Development and Applications. *Personality and Social Psychology Bulletin* 22, 6 (1996), 620–634. https://doi.org/10.1177/0146167296226007 arXiv:0803973233
[17] Mark ter Maat. 2011. How Agents' Turn-Taking Strategies Influence Impressions and Response Behaviors. *Presence: Teleoperators and Virtual Environments* 20, 5 (2011), 412–430. https://doi.org/10.1162/PRES_a_00064