Human, Chameleon or Nodding Dog?*

Virtual Experiments with Non-Verbal Persuasion

Leshao Zhang
Queen Mary University of London
London, UK
leshao.zhang@qmul.ac.uk

Patrick G.T. Healey
Queen Mary University of London
London, UK
p.healey@qmul.ac.uk

ABSTRACT

Immersive virtual environments (IVEs) present rich possibilities for the experimental study of non-verbal communication. Here, the 'digital chameleon' effect, -which suggests that a virtual speaker (agent) is more persuasive if they mimic their addresses head movements-, was tested. Using a specially constructed IVE, we recreate a full-body analogue version of the 'digital chameleon' experiment. The agent's behaviour is manipulated in three conditions 1) Mimic (Chameleon) in which it copies the participant's nodding 2) Playback (Nodding Dog) which uses nods from playback of a previous participant and are therefore unconnected with the content and 3) Original (Human) in which it uses the prerecorded actor's movements. The results do not support the original finding of differences in ratings of agent persuasiveness between conditions. However, motion capture data reveals systematic differences in a) the real-time movements of speakers and listeners b) between the Original, Mimic and Playback conditions. We conclude that the automatic mimicry model is too simplistic and that this paradigm must address the reciprocal dynamics of non-verbal interaction to achieve its full potential.

CCS CONCEPTS

• **Applied computing** → **Psychology**; • **Human-centered computing** → **Virtual reality**; Laboratory experiments; • **Computing methodologies** → **Virtual reality**; **Motion capture**;

KEYWORDS

Digital Chameleons; Behaviour Mimicry; Persuasive; Virtual Reality; Motion Capture; Human Agent Interaction; Non-verbal Communication

ACM Reference Format:
Leshao Zhang and Patrick G.T. Healey. 2018. Human, Chameleon or Nodding Dog?: Virtual Experiments with Non-Verbal Persuasion. In *2018 International Conference on Multimodal Interaction (ICMI '18), October 16–20, 2018, Boulder, CO, USA*. ACM, New York, NY, USA, 9 pages. https://doi.org/10.1145/3242969.3242998

*Nodding Dog: This is a reference to the toy dogs popular in some countries whose head is attached by a spring or wire and bounces in response ambient vibration. In the US they are sometimes referred to as "Bobbleheads" and in China as "摇头狗". The important point is that the apparent head nods are not responsive to the interaction.

ICMI '18, October 16–20, 2018, Boulder, CO, USA

ACM ISBN 978-1-4503-5692-3/18/10.
https://doi.org/10.1145/3242969.3242998

1 INTRODUCTION

Persuasion has naturally attracted attention in studies of non-verbal behaviour because of the potential practical implications for human interaction in business, politics and everyday life. The evidence from social science suggests that, overall, persuasive communicators tend to be more active nonverbally and more responsive to addressees [41, 42]. This includes a range of non-verbal cues including increased eye-contact, greater facial expressiveness, more frequent nods and increased gesturing [18, 23, 41, 43].

The studies that provide evidence for these conclusions face some important methodological challenges. Non-verbal interaction is, of course, complex and dynamic. In natural interaction, a range of different behaviours −facial expressions, gestures, nods, body position, body orientation− are concurrently deployed. It is hard to capture this complexity through detailed observation alone and there is always some uncertainty about how well such observations generalise beyond the cases described [35, 47]. One response to this issue is to generate hypotheses about specific non-verbal behaviours and investigate whether their frequency varies systematically with the rated persuasiveness of a human speakers' performances [18, 43]. This method can reveal systematic patterns but leaves questions of causation open.

Early experimental studies that attempted direct tests of the effects of non-verbal behaviours on persuasion typically tried to manipulate the speaker's performance, e.g. through instructions to be more or less persuasive or through instructions about how to 'perform' persuasiveness such as certain body angle, percentages of gaze and levels of gesturing and smiling [23]. A known problem with this general approach is that 'actors' and confederates find it difficult to control specific aspects of their behaviour in live interaction without also influencing their performance of other tasks and undermining the naturalness of the interaction in general [14, 36]. This problem is substantial, some effects appear to be entirely the result of a confound due to the use of confederates [14].

Immersive Virtual Environments (IVEs) offer a potentially powerful way to address these difficulties because they provide experimenters with greater experimental control of individual behaviours [5]. Rendering or animating behaviours into a virtual environment means that it is possible, in principle, to selectively alter single behaviours while leaving others unchanged. This technique was used in the "Digital Chameleon" study in which an agent covertly mimics a human participant's head movements, at a short delay, while other aspects of message delivery remain constant [9]. This study was designed as a test of Chartrand and Bargh's hypothesis that automatic mimicry tends to increase affiliation [20] and provided evidence that mimicry makes avatars more persuasive [9]. A similar study suggested that mimicry makes avatars more likely to be trusted [49]. However, these results have not been consistently replicated [28, 46].

This paper sets out to investigate this apparent inconsistency in the results. Using a custom built immersive virtual environment (IVE) we replicated the basic head-movement manipulation reported in the Digital Chameleon experiment [9] and add a new experimental condition and analysis. The environment allows participant's full-body immersion with motion capture and selective manipulating agent's head movement enabled by algorithms.

We test the persuasiveness of a virtual agent in three conditions: Mimic ('Chameleon'), Playback ('Nodding Dog') and Original ('Human'). The first two conditions replicate the original Digital Chameleon experiment [9]. Since in these two conditions the agent uses the listener's head movements, we added a third condition involving the original speaker's head movements. The third condition enables a comparison of Mimicry and Playback with original movements of the actors who produced the basic speech used in all three conditions. Naïve participants interact with each agent in the IVE and subsequently rate it on measures of social presence, agreement with and impression of the agent. Participants concurrent movements are captured using optical motion capture.

The results provide no evidence of differences in participant's subjective estimates of: a) social presence b) agreement with the agent c) general impression of the agent in the three different conditions. They also provide no evidence of differences between conditions in participant's overall responsiveness, as indicated by gross measures of their concurrent head movements. However, they do show that the (original) human speakers move much more than their chameleon or nodding dog counterparts and that listener also repeat the human speaker's movements systematically *less* than the chameleon (trivially) or nodding dog (non-trivially). We argue that a key weakness of this paradigm in its current form is the failure to properly address the reciprocal dynamics of natural non-verbal interaction: speakers and hearers in ordinary conversation interact incrementally and selectively to arrive at a joint understanding. Nonetheless, the potential for addressing these more complex dynamics is clear and presents a rich opportunity for the field.

This study replicates the original 'Digital Chameleons' experiment in a new, more realistic behavioral setting; provides a critical 'natural' baseline/control condition not present in the original study; and provides a new, more detailed analysis of the degree of non-verbal co-ordination observed in this setting. The results demonstrate a problem with this paradigm, namely that it misconstrues the basic reciprocal dynamics of natural interaction i.e. speakers do not behave at all like listeners.

2 RELATED WORKS

2.1 Studies of Nonverbal Interaction with Immersive Virtual Environments

An immersive virtual environment (IVE) is one that perceptually surrounds the users, increasing their sense of presence or being within it. It has two features: first, the user's movements are tracked and automatically and continually updated in the virtual environment; second, it enables the construction of a variety of scenarios and tasks. This makes it possible to carry out controlled manipulations of the user's experience [8].

As noted in the introduction, IVEs have many potential advantages compared to other experimental approaches to studying nonverbal interaction. Non-verbal behaviours are often complex and highly inter-correlated. IVE's can provide a high level of selective control over these behaviours [6, 28, 29]. They also give researchers access, in principle, to all of the participant's motion data which can provide useful additional dependent variables for analysis [16].

A critical part of the IVE experimental paradigm is the use of avatars or agents to represent people within the IVE. An avatar is a virtual character directly controlled by a real human whereas an agent is a virtual character driven by algorithms. Avatars create the experience of embodiment for participants which enhances the body-ownership illusion [15, 40]. This can be powerful enough to create experiences, such as people with typical colour vision experiencing being colour blind, and in turn changing their subsequent behaviour toward others [2]. Similarly the cutting a virtual tree in IVE versus reading a text description or watching a video depiction of the tree cutting process is more effecting in encouraging paper conservation [1].

For the purposes of experiments on human interaction, agents can provide a significant advance on the use of confederates producing scripted behaviours [14, 37]. In principle, agent behaviours are perfectly controllable and always blind to the experimental manipulation. This can be exploited to manipulate factors such as personal distance or gaze without changing any other behaviours [7].

2.2 Virtual Realism

A key question for IVE based experiments is whether the avatar/agent and virtual environment are realistic enough to elicit participant's typical responses. The most obvious level at which this issue arises is visual appearance. As noted, there is evidence that virtual bodies are sufficiently realistic to induce the body ownership illusion. Additionally, seeing a realistic virtual body in the same location and posture as the physical body also engenders a sense of body ownership [40]. Latoschik et. al. [38] studied the impact of avatar realism on the experience of embodiment and quality of social interaction. They compared a neutral abstract avatar representation (wooden mannequin) with high fidelity scans of real humans. The results suggested that the realistic avatars were significantly more human-like when used as avatars for the others and more accepted in terms of virtual body ownership.

The issue is not necessarily one of visual realism. Argelaguet [4] found in their hand ownership study that the feeling of agency (sense of controllability of one's own actions) is stronger with less realistic hands which may be due to a trade-off between visual realism and tracking accuracy. This suggests that although a realistic avatar created a higher sense of ownership, it requires fine-grained motion mapping for the feeling of agency. Realism may also be linked to persuasiveness. Guadagno compared the persuasiveness of virtual agent with no nonverbal behaviour with an agent with complex nonverbal behaviour including tracking participants' eye contact. This suggested that the more realistic behaviour is more influential [26]. This points to the importance of motion realism, possibly over visual realism for maintaining the sense of immersed engagement in IVE experiments.

2.3 Interactional Realism

A more critical issue is the level of interactional realism that is achieved and the evidence here is less clear. While there is wide recognition that the non-verbal cues involved in phenomena such as persuasiveness are complex this is usually thought of only from a production point of view - i.e. aspects of the *speaker's* body position, gaze, gesture etc. (see above). However, speaker and addressee's non-verbal behaviours are tightly coupled. Speakers constantly monitor their addressees for concurrent non-verbal feedback and when this absent, mistimed or incongruent, e.g. raised eye-brows or puzzled looks when a smile was expected, speakers adjust their performance -mid turn- to try and get things back on track [13, 24]. Moreover, speakers and listeners behave in systematically different ways e.g. speakers gesture more and addresses gesture systematically less than unaddressed third parties [11, 30, 31]. This highlights the importance of considering the co-ordination of the non-verbal dynamics of conversation across multiple participants i.e. across both speakers and addressees.

Head movements are an especially important conversational signal [12, 17, 27, 32, 33]. They include nods, shakes and changes of angle / orientation, etc. Amongst these head nods are the most frequent behaviours. 81.5% of head movements are head nods when participants were the listeners [51] and 56% when participants were in free conversation [34]. For listeners, head nods serve as: (1) 'back channels' to signal their level of understanding of an ongoing turn, (2) a signal that a (currently unaddressed) listener would like to take the floor, (3) to communicate the degree of understanding, agreement, or support. For speakers, head nods can (1) serve as a signal of the intention to continue speaking, (2) to seek or check agreement, (3) to express emphasis, (4) to control and organize the interaction, (5) as 'beat' gestures that accompany the rhythmic aspects of speech, (6) to signal lexical repairs and (7) mark switches between direct and indirect discourse. Listener's head nods as feedbacks are often concurrent with the speaker's turn. Single head nods or jerks are the most frequent feedback movements. Minimal head nods are used to show continuation of contact, perception and understanding. More complex head movements are used for emphasis, agreement, disagreement, surprise and disappointment. Multiple nods or sequences of expressions, e.g. nods and smile, can be used to acknowledge/refuse an idea or to ask for clarification [3].

2.4 Digital Chameleons?

The term Digital Chameleons was coined by Bailenson and Yee [9] and builds on the ideas of Chartrand and Bargh [20] who claim that people automatically mimic each others' movements and behaviours unconsciously during interaction, usually within a short window of time of between three to five seconds [21]. In addition, it is claimed that mimicry has various effects including changing individuals' cognitive processing style, altering performance on tests of ability, creativity and shifting preferences for consumer products. Importantly, it is also thought to improve liking, empathy, affiliation, increase help behaviour and reduce prejudice between interactants [21]. This leads to the prediction that if one person simply repeats aspects of another person's movements this ought to enhance their credibility and persuasiveness for the other participant. The degree of mimicry must, of course, be carefully controlled to avoid creating a sense of parody or irony. However, nodding is ostensibly a good candidate since it is a relatively unmarked and positively valenced behaviour.

Bailenson and Yee [9] tested this prediction using a virtual agent who delivers a message that aims to convince students that they should always carry their ID card. The message was delivered to seated participants through a Head Mounted Display after which their post-hoc ratings of the persuasiveness of the speech by agent were measured by questionnaire. The experiment compared the persuasiveness of the agent in two different conditions: mimic and non-mimic. In the mimic condition, the agent reproduced exactly the head movements of each participant with a fixed 4s delay. While in the non-mimic condition (corresponding to the Playback condition here), the agent's head movements were controlled by a canned sequence recorded from the head movements of the previous participant. The results provided evidence that the mimicking agents were more persuasive than the non-mimicking agents. Further more, the motion analysis suggested that the overall movements of men were significantly larger than women.

In a second study Bailenson and Yee [10] found a relationship between liking and mimicry using a mechanical device mimicking participants' handshakes and suggests that males responded more strongly to mimicry than females. Verberne et. al. [49] used virtual agents to mimic participants' head movements during a route planner game and investment game. The results showed that the mimicking agents were more liked and trusted by participants in route planner game rather than investment game. Verberne suggested that the Digital Chameleons effect comes with conditions: first, the effect might need a certain time to be obtained; second, the type of behaviour might be a factor, i.e. might depend on the consequence of the behaviour being predictable. Stevens et al. studied the likability of an agent which mimicked participant's head nods or eyebrow raises when the agent repeated a participant's sentence. They suggested that the likability of the agent could depend on mimicry and how it is moderated by the prominence of visual cues. More prominent the cues the participant produced, the greater the judged lifelikeness of the agent in the mimic condition [48].

However, the Digital Chameleons effect has not been consistently replicated. Riek et. al. [46] compared human-robot rapport in three conditions: robot mimics participants' full head movement, robot mimics participants head nodding only and robot does not mimic participants. They found there is no significant difference in human-robot rapport between the three conditions. They suggested that the survey measure might be unsuitable for rating human-robot interaction and also noted that men had larger movements than women and that participants' movements were possibly unexpected. In an especially careful study Hale et. al. [28] extended mimicry to head and torso movements, and measured the rapport and trust towards the agent after participant and agent had carried out a photo description tasks. The results suggested mimicry has no significant relationship with human-agent rapport or trust.

The current study is designed to provide another replication of the original study [9] but also to explore more carefully the interactional context, especially the relationship between speaker's

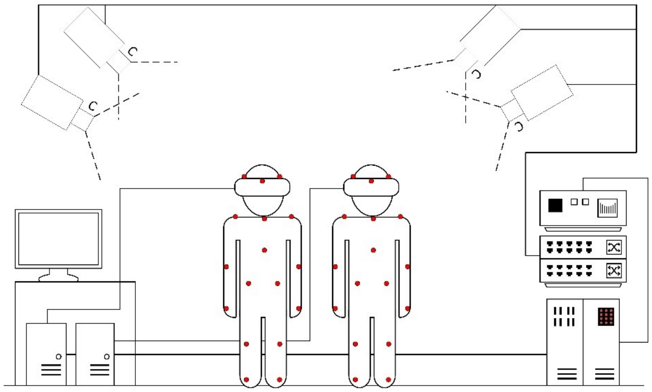


Figure 1: Schematic Diagram of the Capture Set-up in the Human Interaction Lab

and listener's head movements. To do this we recreate the two original conditions and added a third condition.

(1) *Mimic (Chameleon)* – the agent mimics a participant's head movements at a 4s delay.

(2) *Playback (Nodding Dog)* – agent replays the previous participant's head movements.

(3) *Original (Human)*: the agent replays the original speaker's movements.

In the original 'Digital Chameleons' study, the playback condition was designed as a control comparison in which the agent would move as much as in the mimic condition but without any connection to the human participant's ongoing head movements. However, this also entails that the agent's head movements are unconnected with the content. In the context of the literature on human interaction cited above this means that it is likely that they depart significantly from a speaker's natural head movements and are timed in a way that is out of step with the normal dynamics of a conversation exchange. With the third condition, we were also able to compare the persuasiveness of the agents mimicking participant's head movements to the natural speaker's (agents) movements.

3 EXPERIMENT

We built an IVE supporting human-agent interaction that replicates the Digital Chameleon paradigm but in a somewhat more realistic virtual environment and with full-body avatars. The experiment took place in the Human Interaction Lab of Queen Mary University of London. Participants took a first person perspective on their avatar. A Vicon Mocap system was used to provide full body motion capture (Figure 1). Participants wore lycra suits with markers attached and a Head Mounted Display (HMD) in the standing position (Figure 2). The mocap data was synchronized to the virtual avatar so all participant movements could be recreated in the virtual world.

The virtual environment presented an office in the daytime. The agent stood next to the office desk. Male and female virtual agents and male and female avatars were used for participants. Figure 3 shows the virtual environment from the participant's point of view. Participants see their body movements in real-time.

Figure 2: Participants wore black suits attached with markers and Head Mounted Display (HMD)

Figure 3: The virtual environment in the participant's view

One female and one male confederate were recruited to produce the agents behaviour. They performed a scripted speech about the college regulations relating to student ID cards. Their natural body movements were recorded by the motion capture system and used to create the agent animations.

Apart from the body movements, the agent blinked randomly and had slight eye movements. The agent's lip movements were driven by the amplitude of the recording of the confederates.

3.1 Participants

52 participants were recruited by email, posters and through a participant panel. One participant had to be excluded because of problems in data logging. Each participant received 10 pounds or 4 credits for their participation. The final sample consisted of 29 female and 22 male students between 18 to 52 (Mean=21.75,

SD=5.836). None of the participants reported severe motor, auditive or visual disabilities/disorders.

3.2 Procedure

After welcoming the participants to the lab, they were introduced to the purpose of the study and asked to sign a consent form. After putting on the capture suits and the HMD they were given time for free exploration of the IVE. The virtual agent would then welcome them and give a short introduction as a training session. After that, the virtual agent delivered a 2 minute message asking students to always carry their student ID card.

After the agents finished their speech, participants were asked to take off the HMD and fill out an online questionnaire, the same one used in the Digital Chameleons study, which asked them to assess the ID card regulations: "I agree with the plan to implement ID cards", "I think the proposed ID cards are: valuable, workable, needed"; the impression of the virtual presenter: "The presenter was: friendly, likeable, honest, competent, warm, informed, credible, modest, approachable, interesting, trustworthy, sincere"; and the social presence of the virtual presenter: "To what extent you: enjoyed the experience, want to meet him/her again in current situation, feel him/her isolated, want to meet him/her again, feel comfortable with him/her, feel him/her cooperative, feel self-conscious or embarrassed with him/her"; with Likert scale range from 1 strongly disagree to 7 strongly agree.

Once participants completed the questionnaire, they reentered the IVE and repeated the procedure in a different condition. The gender of the agents was switched, and their behaviour randomly assigned to a different experimental condition. They filled out the questionnaire again after the experiment.

3.3 Independent Variables

The experiment is a mixed design with three independent variables: participant's gender (male or female), agent's gender (male or female) and agent's behaviour (mimic or playback or original). Participant's gender is our between-subjects variable. Agent's gender and agent's behaviour are our within-subjects variables.

3.4 Manipulations

Participants were assigned to the mimicry, playback and original conditions in rotation. In the mimic condition, the agent's head nods exactly mimicked those of the participant at a 4-s delay. In the playback condition, the agent's head nods were an exact replay of the nods from the previous participant. This ensured that the agent moved as much in the mimic and playback conditions but with different timing. In the original condition, the agent played a captured movement of an experiment confederate delivering the message. In all the three conditions, only the head pitch of the agent was manipulated, and the yaw and roll of the head was kept as the original condition.

3.5 Measures and Hypotheses

3.5.1 Effectiveness of the Agent. We replicated the composite measure of the agent's effectiveness from the Digital Chameleon study by taking the mean response to the 4 agreement questions (how much the participant agreed with the agent's persuasive message), 13 (12+overall) questions on impressions of the agent, and 8 (7+overall) questions on the agent's social presence. Cronbach's alpha for this composite measure of these 25 items was .854. To provide a conservative test of the 'Digital Chameleons' effect we used a non-directional null hypothesis:

Null Hypothesis 1. *There is no difference in the effectiveness of the agent between the different conditions (mimic or playback or original).*

3.5.2 Head Movements. We recorded participants' head movements at approximately 60 Hz and determined the maximum value (in degrees) of deviation from the straight-ahead position (i.e., looking directly at the agent's eyes). We summed up the participants' total head movements in the 2 minutes talk. Our hypothesis was that head movements would serve as an approximation of visual attention, and that participants would look farther away from the agent in the recorded condition than in the mimic condition because mimicking agents would be more effective at keeping participants' attention. Participants' total head movements would show their listening status, and participants would move less as a listener than a speaker. Similarly, our null hypothesis was:

Null Hypothesis 2. *There is no difference in participant's max or total head movement between the different conditions varies from the participant's gender, agent's gender and agent's behaviour (mimic or playback or original).*

4 RESULTS

4.1 Effectiveness of the Agent

A Shapiro-Wilk test showed that ratings of effectiveness did not deviate significantly from a normal distribution (p=.482). As the experiment has repeated trials, we ran a Generalized Linear Mixed Model (GLMM) analysis with three between-subjects as fixed factors (participant's gender, agent's gender, and agent's behavior) and agent's effectiveness as the target; subject and agent appearance were included as random effects; and normal model was used. The results suggested no significant effect of agent effectiveness on condition ($F_{2,93}$=2.147, p=.123), participant's gender ($F_{1,93}$=.633, p=.428) or agent's gender ($F_{2,93}$=.082, p=.776). Table 1 shows the results for the three underlying measures of the agent's effectiveness, α is the Cronbach's alpha for the composite measure, p value from one-way ANOVA. The results suggested no significant difference was found between conditions for social presence, agreement and impression on the agents.

4.2 Maximum Head Nodding Movement

Following Bailenson we calculated the maximum head nodding movement of each participant. This measure was strongly positively skewed. The log of the raw data followed the normal distribution (p=.189). We ran the GLMM with max head movement as target, log normal model and all other settings same as before. The result found no effect of agent's behavior ($F_{2,93}$=.734, p=.483) but a main effect of participant's gender (Figure 4), with male participants straying farther away from center (Mean = 17.745, SE = 1.727) than female participants (Mean = 11.712, SE = 1.027), $F_{1,93}$=9.016, p<.01. No interactions were significant. No significant correlation between head nodding and agent's effectiveness, r_s=.012, p>.05.

Table 1: Results for the components of the agent's effectiveness measure

Measure	Number of items	α	Mimic condition		Playback condition		Original condition		p
			Mean	SD	Mean	SD	Mean	SD	
Social presence	8	.762	31.05	8.03	34.05	6.96	29.98	6.93	.08
Agreement	4	.891	21.51	4.33	22.21	4.16	22.34	4.47	.69
Impression	13	.926	59.27	14.55	65.84	14.34	62.04	12.94	.15

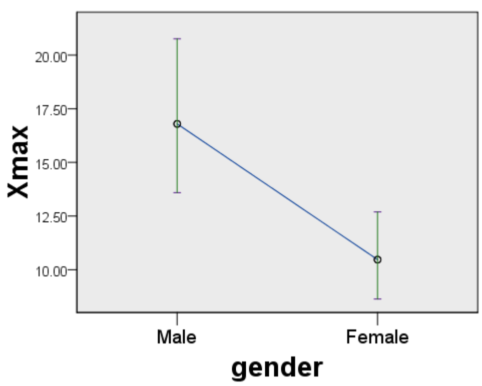


Figure 4: Gender difference in maximum head nodding movement

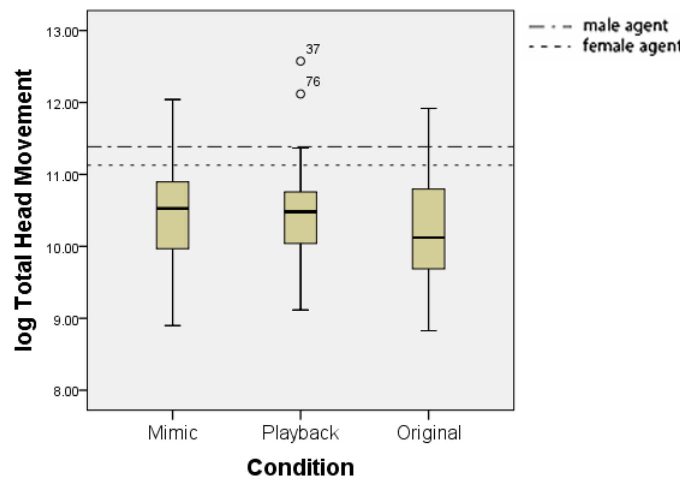


Figure 5: The boxplot of log participants' total head movement. $Mean_m$**=10.483,** SD_m**=.747;** $Mean_p$**=10.46,** SD_p**=.760;** $Mean_o$**=10.299,** SD_o**=.832. The 95% upper confidence interval for mean is 10.74 for mimic and playback condition, 10.59 for original condition. The log total head movement for male agent is 11.3846, for female agent is 11.1274.**

4.3 Total Head Movement

The total head movement also does not fit normal distribution and positive skewed. A Shapiro-Wilk test showed that the log of total head movement did not deviate significantly from a normal distribution (p=.191). The same GLMM was run with total head movement as target. No main effect found on participants' total head movement. The one sample T-test of the log agent's total head movement over the log participants' head movement suggested participants' total head movement was significantly lower than agent's total head movement. For male agent, $t(99) = -12.428, p < .001$; for female agent, $t(99) = -9.132, p < .001$. Figure 5 shows the boxplot of log participants' total head movement. The total head movement of the male agent is represented by the dash dot line and the female agent by the dotted line. From the plot we can see that the total head movement of male agent is higher than female agent and higher than the 95% upper confidence interval for mean of participant's total head movement.

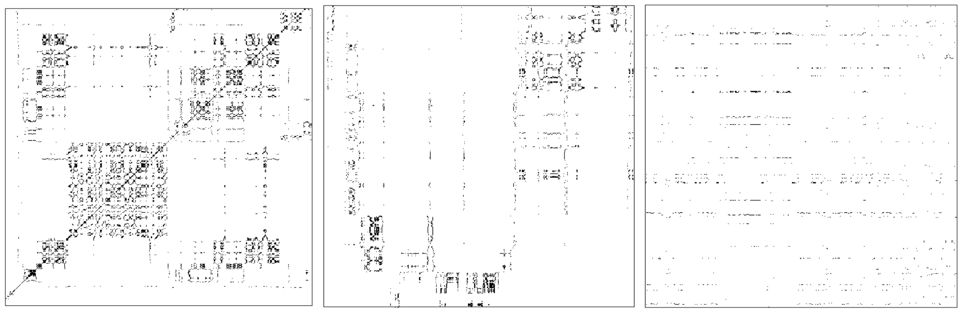
4.4 Cross Recurrence Quantification Analysis

Recurrence Quantification Analysis (RQA) [50] is a time series analysis method used on non-linear dynamic systems. Cross Recurrence Quantification Analysis (xRQA) is RQA applied to two independent time series e.g. two participants in a conversation [22, 44, 45].

RQA produces Recurrence Plots (RP) used to identify the overal patterns of behaviour repetition, and three quantiative measures which can be used to identify the degree of coordination for certain behaviours – %REC is the percentage of recurrence points in the RP. It tells how much the two time series are repeated. %DET is the percentage of recurrence points which form diagonal lines. It shows how much one time series is predictable from another; LMAX is the length of the longest diagonal line segment in the RP, excluding the main diagonal line of identity (i = j). The shorter the LMAX, the more chaotic (less stable) the signal.

We carried out xRQA on the motion capture data. The two independent time series were the participant's head nodding and the agent's head nodding. The time interval for one unit was 15ms. The parameters were chosen based on the instructions of the technique [25, 50]. i.e. Embedding Dimension=10, Delay=1, Radius=50. Figure 6 is an example of head nodding recurrence plot in different condition. As expected mimicry behaviour produces a lot of square recurrence patterns (Figure 6a). This is guaranteed by the experimental manipulation. The recurrence pattern in the playback condition is different with many more of vertical lines (Figure 6b) and the recurrence pattern in the original condition is mostly horizontal line and much more fragmented (Figure 6c).

Three GLMM analyses of the results were carried out for %REC, %DET and LMAX, with three between-subjects fixed factors (participant's gender, agent's gender and agent's behavior), %REC, %DET and LMAX as the target and subject and agent appearance as random factors and an Inverse Gaussian distribution. Figure 7 shows the pattern of main effects and interactions on %REC, %DET and LMAX.



(a) Mimic condition **(b) Playback condition** **(c) Original condition**

Figure 6: Example of recurrence plot in Mimic, Playback and Original conditions

The experimental manipulation in the mimicry condition guarantees that LMAX (≈10000) is greater than the other two conditions (LMAX<200) since it ensures perfect mimicry at a 4 second delay. Ignoring this there is a main effect of agent's behaviour (p<.001) on %REC, %DET and LMAX (Table 2). %REC, %DET and LMAX are significantly higher for playback than for original. Agent's gender was the second main effect on %DET. Female agent was significantly higher than male agent in %DET, $F_{1,54}$=13.352, p<.01. There are also main effects of participant's gender, agent's gender and interaction between agent's behaviour and participant's gender on LMAX. Pairwise contrasts showed that LMAX was significantly higher for male participants than for female participants $F_{1,54}$=9.22, p<.01; LMAX was significantly higher for female participants than for male participants, $F_{1,54}$=5.467, p<.05; For both male ($F_{1,54}$=44.712, p<.001) and female ($F_{1,54}$=76.171, p<.001) participants, LMAX was significantly higher for playback than for original.

5 DISCUSSION

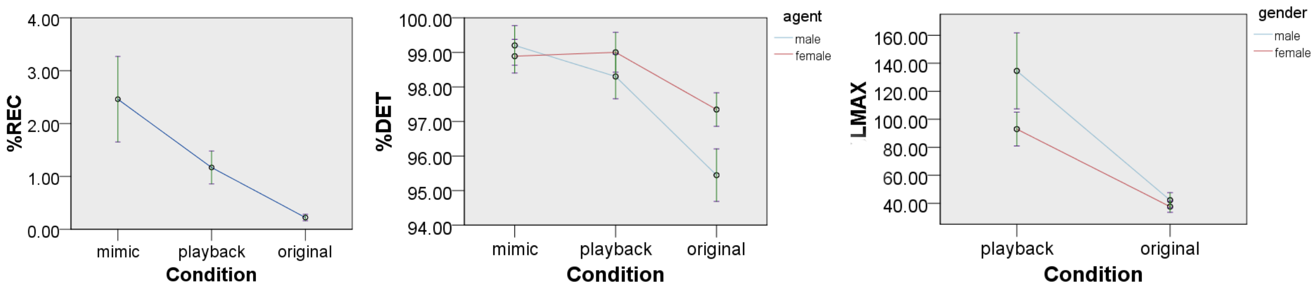
The null results of the effectiveness of the agent suggest that there is a great chance that we have to accept the null hypothesis 1. That is, the differences in agent behaviour have no significant effect on perceived persuasiveness and, as such, fail to replicate the basic 'digital chameleon' effect. Two possible explanations for this are: i) the Digital Chameleons effect is not reliable across different experimental situations; ii) this experiment fails to replicate some critical feature of the original.

Our results broadly favour the first explanation. It highlight some important problems with the automatic mimicry idea and its implementation in the Digital Chameleons study. First, speaker's normally move their heads more frequently and in a richer variety of ways than listeners. If the speaker simply copies a listener's head movements this represents a significant departure from normal behaviour. This might have effects on persuasiveness in certain circumstances but it breaks the balance of initiative typical of ordinary conversation. Secondly, if virtual agents 'blindly' mimic participant's head movements this breaks the relationship between their head movements and the content of their speech. It seems unlikely that this can improve persuasiveness unless by chance it lines up with what is being said. Third, the Digital Chameleons is based on the assumption that the chameleon effect could be produced by an automatic algorithm implemented on the virtual agent. However, this algorithm does not reproduce real human's mimicking behaviour which is more flexible. Research shows that contingency might be more important than similarity in human mimicry behaviour [19]. Although unconscious human mimicry of an interaction partner might increase social influence, it does not entail that an agent would also gain social influence by implementing simple automatic mimicry rules. This highlights a gap in the theory which assumes we don't automatically mimic *everything* - since conversation would grind to a complete halt. However, it is unclear –from the point of view of implementation– how it is moderated in terms of timing, choice of behaviours or with respect to particular social goals.

It is possible that an agent would gain social influence by mimicry if it applied a more selective algorithm to judge when it should perform a mimicry task in the social interaction context. Lee et. al. [39] compared the perception of inappropriate head nods generated by three different method for a virtual agent: by a machine learning data-driven approach, by a handcrafted rule-based approach and by a human. The results suggest a data-driven approach had the best performance over all the three methods.

Nonetheless, the failure to replicate may be due to differences in experimental setting. We used a more complex virtual environment and full body interaction. The justification for this is greater naturalism but it may also bring in some unpredictable factors such as the uncanny valley effect. Agents in the mimic and playback condition might be perceived as weird to the participants given greater expectations about behaviour or simply the greater range of behaviours. This might weaken or dilute the Digital Chameleon effect. Like Verberne et. al. [49] pointed out, this effect might only work with certain behaviours or it might need a long time to take effect. However this line of argument leads to the conclusion that the effect is not robust. Another possibility is that students attitudes to the importance of ID cards have changed. It is notable that the subjective ratings in the survey were relatively high and it is possible that this obscures potential differences in persuasiveness.

The result of participant's max and total head nodding movement also suggested that we can partly reject the null hypothesis 2. That is, although agent's behaviour or agent's gender did not affect participant's head movements, female participants move their heads significantly less than male participants. This gender difference in head movements was also reported in the 'Digital Chameleons'



(a) %REC (b) %DET (c) LMAX

Figure 7: Main effects on %REC, %DET and LMAX. Figure 7a is the effect of agent's behaviour on %REC; figure 7b is the effect of the interaction between agent's behaviour and agent's gender on %DET; figure 7c is the effect of the interaction between agent's behaviour and participant's gender on LMAX, mimic condition excluded.

Table 2: Results of the GLMM analysis of %REC, %DET and LMAX in different agent's behaviour condition

item	Mimic condition		Playback condition		Original condition		F	df1	df2	p
	Mean	SE	Mean	SE	Mean	SE				
%REC	2.461	.407	1.169	.156	.223	.031	32.201	2	85	<.001
%DET	99.046	.189	98.652	.218	96.396	.227	43.837	2	85	<.001
LMAX	10290.478	5103.501	114.129	7.820	40.707	2.581	41.760	2	91	<.001

study. It suggests that female participants are more focused when listening to the speech. On the other hand, the basic comparison between participant's and agent's total head movement shows that participants as listeners make only limited head movements. This is a strong evidence of the first problem addressed above.

Perhaps more interesting are the results of the cross recurrence quantification analysis of head movements. These show some clear differences in behaviour patterns for the interactions in three difference condition. The head movements of the agents and participants had highest levels of repetition (%REC, %DET and LMAX) in the mimicry condition (as expected); medium %REC, higher %DET and lower LMAX in playback condition; lowest %REC, %DET and LMAX in original condition.

Since in the mimicry condition, the agent repeats the participant's head movements with a 4s delay it has the highest %REC, %DET and LMAX. More importantly, the playback condition has higher %REC, %DET and LMAX than in the original condition. In the playback condition, the xRQA correlates two separate participants' head movements, and as listener head movements were generally reduced this can lead to higher repetition of (non) movement. In contrast to this in the original condition, the subjects of xRQA were the head movements of a participant as a listener and the head movements of a confederate as a speaker. As noted, speakers move their heads significantly more than listeners and this means the chance level of matching non-moving heads is much lower. The implication of this is that natural interaction is actually characterized by *low* levels of speaker-listener mimicry because of asymmetries in the both the level of head movements and their functions. This is consistent with the literature on human interaction but incompatible with an automatic mimicry model.

6 CONCLUSIONS

This study attempted to replicate the Digital Chameleon effect using a realistic IVE and virtual characters with motion captured animations. The implemented motion capture system provided participants with full-body avatars in the IVE in a relatively realistic environment but the results do not support the claim that a mimicking agent is more persuasive than a human narrator or a nodding dog. Analysis of motion capture data suggests the presence of gender difference when participants interact with the agents. Furthermore, cross recurrence quantification analysis on the captured motion data reveals differences in coordination of interactions in the mimic, playback, and original conditions. There is a limitation of our study. Although we used a third condition – original intended to test the persuasiveness of a natural interaction, it is only a prerecorded animation. A more conclusive test of the importance of mimicry requires manipulation of live interactions.

ACKNOWLEDGMENTS

The work is supported by EPSRC and AHRC Centre for Doctoral Training in Media and Arts Technology (EP/L01632X/1).

REFERENCES

[1] Sun Joo Ahn, Jeremy N. Bailenson, and Dooyeon Park. 2014. Short- and long-term effects of embodied experiences in immersive virtual environments on environmental locus of control and behavior. *Computers in Human Behavior* 39, February (2014), 235–245. https://doi.org/10.1016/j.chb.2014.07.025

[2] Sun Joo (Grace) Ahn, Amanda Minh Tran Le, and Jeremy Bailenson. 2013. The Effect of Embodied Experiences on Self-Other Merging, Attitude, and Helping Behavior. *Media Psychology* 785022367 (2013). https://doi.org/10.1080/15213260701291346

[3] Jens Allwood and Loredana Cerrato. 2003. A study of gestural feedback expressions. In *First nordic symposium on multimodal communication*. Copenhagen, 7–22.

[4] Ferran Argelaguet. 2016. The Role of Interaction in Virtual Embodiment : Effects of the Virtual Hand Representation The Role of Interaction in Virtual Embodiment : Effects of the Virtual Hand Representation. , 3 – 10 pages.

[5] Jeremy N Bailenson, Jim Blascovich, Andrew C Beall, and Jack M Loomis. 2001. Equilibrium theory revisited: Mutual gaze and personal space in virtual environments. *Presence: Teleoperators & Virtual Environments* 10, 6 (2001), 583–598.

[6] Jeremy N. Bailenson, Jim Blascovich, Andrew C. Beall, and Jack M. Loomis. 2001. Equilibrium Theory Revisited: Mutual Gaze and Personal Space in Virtual Environments. *Presence: Teleoperators and Virtual Environments* 10, 6 (2001), 583–598. https://doi.org/10.1162/105474601753272844 arXiv:arXiv:1011.1669v3

[7] Jeremy N. Bailenson, Jim Blascovich, Andrew C. Beall, and Jack M. Loomis. 2003. Interpersonal Distances in Virtual Environments. *Personality and Social Psychology Bulletin* 29, 7 (2003), 819–833. https://doi.org/10.1177/0146167203253270

[8] Jeremy N. Bailenson and Nick Yee. 2003. The Use of Immersive Virtual Reality in the Learning Sciences: Digital Transformations of Teachers, Students, and Social Context. *Journal of the Learning Sciences* 12, 3 (2003), 361–407. https://doi.org/10.1207/S15327809JLS1203

[9] Jeremy N Bailenson and Nick Yee. 2005. Digital Chameleons. *Society* 16, 10 (2005), 814–819. https://doi.org/10.1111/j.1467-9280.2005.01619.x

[10] Jeremy N. Bailenson and Nick Yee. 2007. Virtual interpersonal touch and digital chameleons. *Journal of Nonverbal Behavior* 31, 4 (2007), 225–242. https://doi.org/10.1007/s10919-007-0034-6

[11] Stuart A Battersby and Patrick G T Healey. 2010. Head and Hand Movements in the Orchestration of Dialogue. *Proceedings of the Cognitive Science Society* 32, 32 (jan 2010). http://escholarship.org/uc/item/4v80m010

[12] Stuart A Battersby and Patrick G T Healey. 2010. Using head movement to detect listener responses during multi-party dialogue. In *Proceedings of LREC workshop on multimodal corpora advances in capturing coding and analysing multimodality.* 11–15.

[13] Janet Beavin Bavelas, Linda Coates, and Trudy Johnson. 2000. Listeners as co-narrators. *Journal of personality and social psychology* 79, 6 (2000), 941.

[14] Janet Beavin Bavelas and Sara Healing. 2013. Reconciling the effects of mutual visibility on gesturing: A review. *Gesture* 13, 1 (2013), 63–92.

[15] Ilias Bergström, Konstantina Kilteni, and Mel Slater. 2016. First-person perspective virtual body posture influences stress: a virtual reality body ownership study. *PloS one* 11, 2 (2016), e0148060.

[16] Jim Blascovich, Jack Loomis, Andrew C Beall, Kimberly R Swinth, Crystal L Hoyt, and Jeremy N Bailenson. 2002. Immersive Virtual Environment Technology as a Methodological Tool for Social Psychology. 13, 2 (2002), 103–124.

[17] Max Boholm and Jens Allwood. 2010. Repeated head movements, their function and relation to speech. In *Proceedings of LREC workshop on multimodal corpora advances in capturing coding and analysing multimodality.* Citeseer, 6–10.

[18] Judee K Burgoon, Thomas Birk, and Michael Pfau. 1990. Nonverbal behaviors, persuasion, and credibility. *Human communication research* 17, 1 (1990), 140–169.

[19] Caroline Catmur and Cecilia Heyes. 2013. Is it what you do, or when you do it? The roles of contingency and similarity in pro-social effects of imitation. January (2013), 1–25.

[20] Tanya L. Chartrand and John A. Bargh. 1999. The chameleon effect: The perception–behavior link and social interaction. *Journal of Personality and Social Psychology* 76, 6 (1999), 893–910. https://doi.org/10.1037/0022-3514.76.6.893

[21] Tanya L Chartrand and Jessica L Lakin. 2013. The antecedents and consequences of human behavioral mimicry. *Annual review of psychology* 64 (2013), 285–308.

[22] Rick Dale and Michael J Spivey. 2006. Unraveling the dyad: Using recurrence analysis to explore patterns of syntactic coordination between children and caregivers in conversation. *Language Learning* 56, 3 (2006), 391–430.

[23] Joyce A Edinger and Miles L Patterson. 1983. Nonverbal involvement and social control. *Psychological Bulletin* 93, 1 (1983), 30.

[24] Charles Goodwin. 1979. The Interactive Construction of a Sentence in Natural Conversation. (1979), 97–121. http://www.sscnet.ucla.edu/clic/cgoodwin/79inter{_}sent.pdf

[25] Allison Gray. 2016. Cross Recurrence Quantification Analysis Tutorial. Retrieved 22 April, 2018 from https://quantdev.ssri.psu.edu/tutorials/cross-recurrence-quantification-analysis

[26] Rosanna E Guadagno, Jim Blascovich, Jeremy N Bailenson, and Cade McCall. 2007. Virtual humans and persuasion: The effects of agency and behavioral realism. *Media Psychology* 10, 1 (2007), 1–22. https://doi.org/10.108/15213260701300865

[27] Uri Hadar, Timothy J Steiner, and F Clifford Rose. 1985. Head movement during listening turns in conversation. *Journal of Nonverbal Behavior* 9, 4 (1985), 214–228.

[28] Joanna Hale and Antonia F. De C. Hamilton. 2016. Testing the relationship between mimicry, trust and rapport in virtual reality conversations. *Scientific Reports* 6 (2016). https://doi.org/10.1038/srep35295

[29] Patrick G. T. Healey, Christopher Frauenberger, Marco Gillies, and Stuart Battersby. 2009. Experimenting with Non-Verbal Interaction. (2009), 2–3.

[30] Patrick G T Healey, Mary Lavelle, Christine Howes, Stuart Battersby, and Rosemarie McCabe. 2013. How listeners respond to speaker's troubles. In *Proceedings of the Annual Meeting of the Cognitive Science Society*, Vol. 35.

[31] Patrick G T Healey, Nicola Plant, Christine Howes, and Mary Lavelle. 2015. When words fail: collaborative gestures during clarification dialogues. In *2015 AAAI Spring Symposium Series,(Chicago).*

[32] Dirk Heylen. 2005. Challenges ahead: Head movements and other social acts in conversations. In *Proceedings of Joint Symposium on Virtual Social Agents.* 45–52.

[33] Dirk Heylen. 2006. Head gestures, gaze and the principles of conversational structure. *International Journal of Humanoid Robotics* 3, 03 (2006), 241–267.

[34] Carlos Toshinori Ishi, Hiroshi Ishiguro, and Norihiro Hagita. 2014. Analysis of relationship between head motion events and speech in dialogue conversations. *Speech Communication* 57 (2014), 233–243.

[35] Adam Kendon. 1990. *Conducting interaction: Patterns of behavior in focused encounters.* Vol. 7. CUP Archive.

[36] Anna K Kuhlen and Susan E Brennan. 2013. Language in dialogue: when confederates might be hazardous to your data. *Psychonomic bulletin & review* 20, 1 (2013), 54–72.

[37] Anna K. Kuhlen and Susan E. Brennan. 2013. Language in dialogue: when confederates might be hazardous to your data. *Psychonomic Bulletin & Review* 20, 1 (2013), 54–72. https://doi.org/10.3758/s13423-012-0341-8

[38] Marc Erich Latoschik, Daniel Roth, Dominik Gall, Jascha Achenbach, Thomas Waltemate, and Mario Botsch. 2017. The effect of avatar realism in immersive social virtual realities. In *Proceedings of the 23rd ACM Symposium on Virtual Reality Software and Technology.* ACM, 39.

[39] Jina Lee, Zhiyang Wang, and Stacy Marsella. 2010. Evaluating models of speaker head nods for virtual agents. In *Proceedings of the 9th International Conference on Autonomous Agents and Multiagent Systems: volume 1-Volume 1.* International Foundation for Autonomous Agents and Multiagent Systems, 1257–1264.

[40] Antonella Maselli and Mel Slater. 2013. The building blocks of the full body ownership illusion. *Frontiers in human neuroscience* 7 (2013), 83.

[41] Marianne Schmid Mast and Gaëtan Cousin. 2013. Power, Dominance, and Persuasion. In *Nonverbal Communication (Handbooks of Communication Science, HOCS 2)*, Judith A Hall and Mark L Knapp (Eds.). DE GRUYTER, Berlin, Boston, Chapter 20, 613–635. https://doi.org/10.1515/9783110238150

[42] Albert Mehrabian. 2017. *Nonverbal communication.* Routledge.

[43] Albert Mehrabian and Martin Williams. 1969. Nonverbal concomitants of perceived and intended persuasiveness. *Journal of Personality and Social psychology* 13, 1 (1969), 37.

[44] Daniel C Richardson and Rick Dale. 2005. Looking to understand: The coupling between speakers' and listeners' eye movements and its relationship to discourse comprehension. *Cognitive science* 29, 6 (2005), 1045–1060.

[45] Michael J Richardson, Stacy Lopresti-Goodman, Marisa Mancini, Bruce Kay, and RC Schmidt. 2008. Comparing the attractor strength of intra-and interpersonal interlimb coordination using cross-recurrence analysis. *Neuroscience Letters* 438, 3 (2008), 340–345.

[46] Laurel D Riek, Philip C Paul, and Peter Robinson. 2010. When my robot smiles at me: Enabling human-robot rapport via real-time head gesture mimicry. *Journal on Multimodal User Interfaces* 3, 1-2 (2010), 99–108.

[47] Albert E Scheflen. 1964. The significance of posture in communication systems. *Psychiatry* 27, 4 (1964), 316–331.

[48] Catherine J Stevens, Bronwyn Pinchbeck, Trent Lewis, Martin Luerssen, Darius Pfitzner, David MW Powers, Arman Abrahamyan, Yvonne Leung, and Guillaume Gibert. 2016. Mimicry and expressiveness of an ECA in human-agent interaction: familiarity breeds content! *Computational cognitive science* 2, 1 (2016), 1.

[49] Frank M F Verberne, Jaap Ham, Aditya Ponnada, and Cees J H Midden. 2013. Trusting Digital Chameleons : The Effect of Mimicry. (2013), 234–245.

[50] Charles L Webber Jr and Joseph P Zbilut. 2005. Recurrence quantification analysis of nonlinear dynamical systems. *Tutorials in contemporary nonlinear methods for the behavioral sciences* (2005), 26–94.

[51] Marcin Włodarczak, Hendrik Buschmeier, Zofia Malisz, Stefan Kopp, and Petra Wagner. 2012. Listener head gestures and verbal feedback expressions in a distraction task. In *Feedback Behaviors in Dialog.*