Analyzing First Impressions of Warmth and Competence from Observable Nonverbal Cues in Expert-Novice Interactions

Beatrice Biancardi
CNRS-ISIR, Pierre and Marie Curie University
Paris, France
biancardi@isir.upmc.fr

Angelo Cafaro
CNRS-ISIR, Pierre and Marie Curie University
Paris, France
cafaro@isir.upmc.fr

Catherine Pelachaud
CNRS-ISIR, Pierre and Marie Curie University
Paris, France
pelachaud@isir.upmc.fr

ABSTRACT

In this paper we present an analysis from a corpus of dyadic expert-novice knowledge sharing interactions. The analysis aims at investigating the relationship between observed non-verbal cues and first impressions formation of warmth and competence. We first obtained both discrete and continuous annotations of our data. Discrete descriptors include non-verbal cues such as type of gestures, arms rest poses, head movements and smiles. Continuous descriptors concern annotators' judgments of the expert's warmth and competence during the observed interaction with the novice. Then we computed Odds Ratios between those descriptors. Results highlight the role of smiling in warmth and competence impressions. Smiling is associated with increased levels of warmth and decreasing competence. It also affects the impact of others non-verbal cues (e.g. self-adaptors gestures) on warmth and competence. Moreover, our findings provide interesting insights about the role of rest poses, that are associated with decreased levels of warmth and competence impressions.

CCS CONCEPTS

• **General and reference** → **Empirical studies**; • **Applied computing** → *Psychology*; • **Human-centered computing** → *Natural language interfaces*;

KEYWORDS

First impressions; Social cognition; Warmth and Competence; Non-verbal behaviour; Human-human interaction.

ACM Reference Format:

Beatrice Biancardi, Angelo Cafaro, and Catherine Pelachaud. 2017. Analyzing First Impressions of Warmth and Competence from Observable Nonverbal Cues in Expert-Novice Interactions. In *Proceedings of 19th ACM International Conference on Multimodal Interaction (ICMI'17).* ACM, New York, NY, USA, 9 pages. https://doi.org/10.1145/3136755.3136779

1 INTRODUCTION

Interaction is a fundamental need of human beings. In everyday life there are many occasions to interact with different people ranging from strangers to very intimates, such as a partner or a member of the family. When we meet strangers, in zero-acquaintance situations such as job interviews or dating encounters, first impressions matter [2].

The first moments of every new encounter are critical, since we often form impressions about others that can have important consequences such as commitment to meet in further encounters, success at job interviews or dating again a potential partner. Goffman et al. [23] define *impression formation* as the process of information perception, organization and integration in order to form coherent impressions of others (e.g., in terms of personality and interpersonal attitudes). We, as people, are aware of these mechanisms and we often attempt to control the impression that others form of us. This latter process is called *impression management* [23], which mainly concerns the control of visual appearance (e.g. hair style, clothing). However, we also attempt to control social behaviour, but it may be difficult to have full control over all social cues that are exhibited during the interaction. In particular, non-verbal behaviours are crucial because they can reveal with high accuracy a variety of information about us including, for instance, sexual orientation [2], personality and interpersonal attitudes [38]. Two dimensions widely impacted by non-verbal behaviours are warmth and competence (W&C) [20], described in section 2.2.

During the last decades, humans have been increasingly exposed to technology, often by interacting with anthropomorphic interfaces, such as humanoid robots or virtual characters. We are interested in first impressions generated by embodied conversational agents (ECAs), which are anthropomorphic virtual characters capable of interaction with user using gestures, facial expression and speech (for more details, see [13]). First impressions in this context are critical, since they affect user's engagement and willingness to continue the interaction [12]. By managing non-verbal behaviours exhibited by a virtual agent we may improve its first impression of W&C on the user.

In order to place this work into a broader perspective, in IMPRESSIONS project, we aim at building an ECA able to manage the first impression user may have, and to adapt its behaviour according to the user's reaction. In our first steps of investigation, in this paper we are focusing on first impressions of W&C in human-human interaction. We start from the analysis of a corpus of mediated human-human interaction, aiming at finding non-verbal cues eliciting different degrees of W&C, with the future purpose of applying these findings in human-agent interaction.

ICMI'17, November 13–17, 2017, Glasgow, UK

ACM ISBN 978-1-4503-5543-8/17/11…$15.00
https://doi.org/10.1145/3136755.3136779

2 BACKGROUND

2.1 First Impressions

First impressions are a psychological process studied from the 40's [3, 4]. More recent models of impression formation have been developed, among them one of the most prominent is the continuum model by Fiske and Neuberg [22]. A common aspect of those models is that forming an impression about a stranger can be described in three steps. First, we *perceive* the person newly met, often visually, as vision is the fastest sensory channel. People immediately perceive and collect information about invariants traits such aS age, sex, ethnicity, and variant traits, such as face, gestures, body posture, gaze. After acquiring this first information, people make *inference* about the other, for example about his/her personality. Finally, the new person is *categorized* in a certain group or subgroup in which the perceiver either feels to fit in or not. Stereotypes often influence impression formation processes [21]. In the next paragraph we describe the nature of the content processed early in impression formation.

2.2 The Two Fundamental Dimensions of Social Cognition

During social interactions many cognitive mechanisms are involved, such as as processing, storing and applying information about other people. These activities are defined as social cognition. Under an evolutionary point of view [20], social cognition reflects the survival need of knowing the intentions of the others (positive or negative, i.e. the warmth dimension), and the consequent ability to enact those intentions (i.e. the competence dimension). Here we chose to adopt the terms of **warmth** and **competence** since they are the most popular: the former includes traits like friendliness, trustworthiness, sociability; the latter includes traits like intelligence, agency and efficacy.

According to the evolutionary explanation, warmth is judged before competence, as others' intentions matter more to survival whether the other can act on those goals. Primacy of warmth is supported by a large evidence [40, 41]. In [41] participants were asked to list the most important personality traits: they listed significantly more warmth traits than competence traits, and the five most frequently listed traits were warmth-related. Moreover, evaluations based on warmth information were strong and stable, while those based on competence information were weak and dependent on accompanying warmth information. Finally, cognitive performance is better for warmth than for competence. For example, in rapidly judging faces at 100ms exposure times, social perceivers judged trustworthiness (as a warmth trait [20]) most reliably, followed by competence [40].

Many other researchers have investigated the fundamental dimensions processed in social cognition, at both person-perception [38] and group stereotype [20] levels. Asch's [4] was the first who intuited the centrality of W&C dimensions in impression formation. Later, Rosenberg [38] distinguished intellectual good/bad traits (such as intelligent, skillful, determined, foolish, unintelligent, irresponsible) and social good/bad traits (such as sociable, honest, warm, unsociable, cold, unhappy) as the main dimensions of person judgments. Wojciszke [42] showed that W&C account for almost 82% of the variance in global impressions of well-known others: when people interpret behaviours or their impressions of others, W&C form basic dimensions that almost entirely account for how people characterize others. Other terms are used in literature to refer to warmth vs competence (other-profitable vs self-profitable traits [35], social vs task orientation [7], communion vs agency [6]), which have been found to overlap in meaning [1].

W&C also underlie group stereotypes, which are formed by combining high versus low levels of these two dimensions. According to Stereotype Content Model [21], competition predicts warmth and status predicts competence, thus social groups are perceived as warm if they do not compete with the in-group for the same resources and they are considered competent if they are high in status (e.g., economically or educationally successful). The role of W&C is crucial because judgments about them elicit unique emotional (admiration, contempt, envy, and pity) [21] and behavioural responses (active and passive, facilitative and harmful) [16].

Many researchers focused on how W&C are mutually perceived, without reaching an overall agreement. Evidence was found for a positive correlation between W&C, called *halo effect* [38]: when describing a person by giving information about only one dimension (warmth or competence), judgments about the other (non-described) dimension tend to go towards the same direction of the described one. However, other studies supported a negative relation between W&C, called *compensation effect* [44]. Judd et al. [24] presented two groups to participants by giving information about behaviours of the group members. Some participants were provided by information about competence (high in one group and low in the other one), others by informations about warmth (again high in one group and low in the other). Participants then judged the two groups both on the dimension that was manipulated and on the other, unmanipulated, dimension. On the manipulated dimension, naturally, the high group was judged higher than the low group. On the unmanipulated dimension the high group was judged lower than the low group.

The co-occurrence of these two opposite effects can be explained by the fact that, while the *halo effect* occurs when one target is judged, the *compensation effect* occurs in two-target comparative contexts [24]. This applies only to W&C and not to other variables, such as healthiness, for which only *halo effect* can occur[44].

Recently, these explanations have been questioned because new findings showed the occurrence of compensation effect also in absence of an explicit comparative context, that is without evoking any explicitly comparison to another target (Kervyn et al. [26] called it *amplification effect*).

In all the studies cited above, W&C impressions were induced using traits or behavioural information, always in a written format. We aim at investigating whether these relations between W&C apply also when impressions are induced by non-verbal cues and in social interactions.

2.3 Non Verbal Cues for Warmth and Competence

Bayes [9] attempted to define and specify the behavioural cues of warmth, by searching for an association between global ratings of warmth and objective measures of specific behavioural cues.

Those included posture, head movements, hand movements, facial expressions and smiling. This last cue was found to be the best single predictor of warmth.

Cuddy et al. [16] studied non-verbal cues conveying W&C. For warmth, they cite the role of Duchenne's smile [18], the presence of immediacy cues that indicate positive interest or engagement (e.g., leaning forward, nodding, orienting the body toward the other), touching and postural openness, mirroring (i.e., copying the non-verbal behaviours of the interaction partner). For coldness, they cite tense posture, leaning backwards, orientating the body away from the other, tense and intrusive hand gestures (e.g., pointing). Concerning competence, they cite non-verbal behaviour related to dominance and power, such as expansive (i.e., taking up more space) and open (i.e., keeping limbs open and not touching the torso) postures. People who express high-power or assertive non-verbal behaviours are perceived as more skillful, capable, and competent than people expressing low-power or passive non-verbal behaviours.

An interesting study on the effect of hand gestures on social perception [27] showed significant effects of hand gestures type on competence perception. In particular, ideationals and object-adaptors (see Section 3.1.2) result in a higher level of competence judgments, compared to absence of gestures, while self-adaptors result in a lower competence. No significant effect of hand gestures was found for warmth.

2.4 Warmth and Competence in Virtual Agents

When building ECAs, social believability is of utmost importance. In particular, exhibiting the appropriate non-verbal behaviours during the interaction with the user is crucial. For this reason, during last decades researchers have increasingly attempted to endow virtual agents with expressive nonverbal behaviour for emotional states [36], personality traits [30] and interpersonal attitudes [37].

More recently, some studies also focused on W&C, we briefly review them in this section. Niewiadomski [34] found that judgments about believability and W&C of an agent are perceived as higher when using multimodal behaviours than one modality alone (either verbal or non-verbal). Moreover, judgments of believability are positively correlated to W&C ones, with the highest effect size for warmth judgments (that is consistent with the idea of a primacy of warmth judgments over competence). These findings highlight W&C's influence on agent's believability and support the hypothesis that, regarding social cognition, people use the same pattern while judging virtual agents and humans.

Bergmann et al. [10] investigated W&C perception in different conditions combining agent appearance and presence/absence of co-speech gestures. They also investigated whether first impressions could change after a second and a longer experience. They found that human-like vs. robot-like appearance provides more stable impressions of warmth, while gestures increase competence ratings.

Nguyen et al. [33] were the first to develop a computational model for W&C, using *"an iterative design methodology tuning the design using theory from theater, animation and psychology, expert reviews, user testing and feedback"*. Videos of actors performing combinations of different degrees of W&C were analyzed by experts in terms of gestures, use of space and gaze behaviours, in order to extract a set of rules to be encoded in a virtual agent. This process was repeated until unanimous satisfaction. An evaluation test showed that users accurately recognized the intended dimensions.

With respect to the previous studies, we still aim at investigating the nature of W&C impressions in human-agent interaction, with a focus on the relations between the two dimensions. With respect to [10], we consider more behaviours than only co-speech gestures, such as facial expressions, trunk leaning and head poses. To find what these behaviours are, we propose a methodology that uses natural interactions videos (see later) with both discrete and continuous annotations.

3 METHODOLOGY

Since our purpose was to investigate W&C perception by analyzing natural human-human interactions, we exploited a corpus of dyadic expert-novice knowledge sharing interactions that is publicly available at https://noxi.aria-agent.eu/. The expert participant was presumed to be knowledgeable about one or more topics that were of interest for the novice. We analyzed the videos of the "expert", since this role is more related to competence expressions, and experts were those who talked more during the dyadic interaction. As we focus on first impressions, we considered the first 5 minutes of the interaction. Moreover, as a first step, we decided to study the perception of W&C from non-verbal behaviour by excluding speech content. Therefore we focused only on the visual modality, leaving aside speech content and prosody features.

3.1 Annotations

For each annotator, we discarded the first annotated video in order to prevent any bias due to the lack of experience of the annotators with the annotation tool. In total, 14 videos, for a total of 70 minutes, were annotated with experts talking about a variety of topics (e.g. travels, video games, cooking recipes). In order to annotate the videos, we used the (Non)Verbal Annotator (NOVA) tool [8], which supports both discrete and continuous annotations.

3.1.1 Continuous Annotations. Continuous annotations were provided by two annotators about their perceived degree of warmth and competence expressed by the expert. Each dimension was separately annotated in a different time.

The difficulty of obtaining consistent annotations of affective content is a well-known challenge. Following Metallinou and Narayanan [32], we adopted some counter-measures: (1) the annotators were motivated and experienced people, with previous experience in affective annotation and background on literature about W&C; (2) since in literature about social cognition (see Section 2.2) W&C are usually described by using a list of traits, instead of providing a unique definition, we adopted the same approach when giving instructions to the annotators about their task, in order to make the task as clear as possible; (3) we discarded the first annotated video in order to prevent any bias due to the lack of experience of the annotators with the annotation tool; (4) we took into account the reaction lag (see Section 4.1); (5) we considered the relative agreement between the annotators (see Section 4.1).

Therefore, for warmth the annotator was asked to evaluate how the speaker seemed "*kind, pleasant, friendly, warm towards his interlocutor*". For competence, its meaning varies according to the context of application. We can consider cognitive competence (knowledge, abstract intelligence and experience), functional competence (skills, accuracy and speed in performing a task) and social competence (relational and behavioural skills, the ability of managing an interaction). Functional competence is not appropriate for our context, because in the database the expert is not performing a practical task. For the remaining two types, we chose cognitive competence, because if the expert is judged on his/her "expertise" about a topic s/he is talking about, this type of information could be useful when modeling the agent's behaviour in the context of the IMPRESSIONS project, where a virtual character has the role of a guide in a museum and the user is likely to form W&C impressions of the agent. Moreover, social competence is related to sociability, one of the main traits representing warmth. By using cognitive competence we can clearly distinguish the two dimensions (W&C) and prevent the annotation from misunderstandings. Thus, the annotator was asked to evaluate how much the speaker seemed "*knowledgeable and expert about the topic he's talking about*".

All the annotations were performed through a tool in NOVA similar to GTrace [15] (see an example of the interface in Figure 1). NOVA allowed the annotators to perform live continuous annotations while watching the videos. Scores ranged from 0 (very low degree of perceived warmth or competence) to 1 (very high degree of warmth or competence), at a sampling of 25 scores per second. Audio of the videos was switched off when annotating as explained above.

3.1.2 Discrete Annotations. Discrete annotations were done at two different times, at a distance of few months, by a single annotator. A high level of agreement between the two sessions was found (Cohen's Kappa >0.6 for each video, 29% of which >0.8, indicating almost perfect agreement). The discrete annotations, described in the following sections, were types of gestures, rest positions, smiling and head movements.

Types of gestures. We combined the taxonomies proposed by McNeill et al. [29] and Bonaiuto et al. [11], and we categorized gestures in 3 main groups. Table 1 summarizes our classification.

Table 1: The gestures categories used in our discrete annotations and their definitions.

Label	Description
beats	Simple, repetitive, rhythmic movements that bear no obvious relation to the semantic content of the accompanying speech.
ideationals	Non-repetitive complex gestures related to the semantic content of the speech.
adaptors	Manipulations either of the person or of some object gestures; often they may serve as the basis for dispositional inferences (e.g., that the speaker is nervous, uncomfortable).

Beats and ideationals gestures are highly related to verbal expression, thus they are made only by the speaker. The difference between the two categories is that ideationals are related to the semantic content of the speech, while beats are less directly so. In addition, ideationals are non-repetitive, more complex and variable in shape than beats, and they often have greater amplitude.

According to McNeill et al. [29], ideationals include:

- *Iconics*: they display, in their form and manner of execution, concrete aspects of the same scene that speech is also presenting. They draw their communicative strength from being perceptually similar to the phenomenon that is being talked about.
- *Metaphorics*: they are similar to iconic gestures in that they make reference to a visual image. However, the images to which they refer pertain to abstractions.
- *Deictics*: they point to a location in the gesture space.

These three subcategories are not easy to distinguish when annotating without audio, since they depend to speech content, pitch and prosody. For this reason we merged them in the ideationals category during our analyses.

Adaptors are not connected to the speech, thus they can occur at any time of the conversation and can be made by both the listener and the speaker. Examples of different types of gestures are showed in Figure 2.

Arm Rest Poses. We can infer important information about others also when their are not performing gestures. Rest position and posture have been found to be possible indicators of communicator's status and attitude [31]. When expert did not perform any gesture (both while speaking and listening to his interlocutor), his rest poses were annotated. We focused on arms' position during the rest pose. All the poses occurring in at least 2 videos are listed in Table 2 and an example of each pose is showed in Figure 3.

Head Movements. We annotated *nods*, vertical up-and-down movements of the head rhythmically raised and lowered, *shakes*, rotation of the head horizontally from side-to-side [25], and *tilts* when the expert's head tilted aside.

4 DATA ANALYSIS

4.1 Pre-processing

The full pre-processing pipeline from raw continuous annotated data to final samples used for the analysis is depicted in Figure 4.

One of the main issues of continuous annotations is reaction lag. In our context, this is the delay between the moment the impression is formed by the annotator and the motor process leading to the concrete annotation made using the mouse using NOVA [28]. We addressed this issue by shifting back 2 seconds the annotations, as recommended by Mariooryad and Busso [28].

The second step of pre-processing was a data smoothing using a simple moving average technique, in order to reduce meaningless noise. When looking for agreement between two or more raters, it is recommended to consider it in relative terms rather than in absolute terms, because of each person's internal scale when assessing affective content [32, 43]. Therefore, before comparing the annotations of the two raters, we discretized the continuous annotations by following the approach proposed by Cowieand McKeown [15] and applied by Chollet et al. [14], considering the relative agreement of

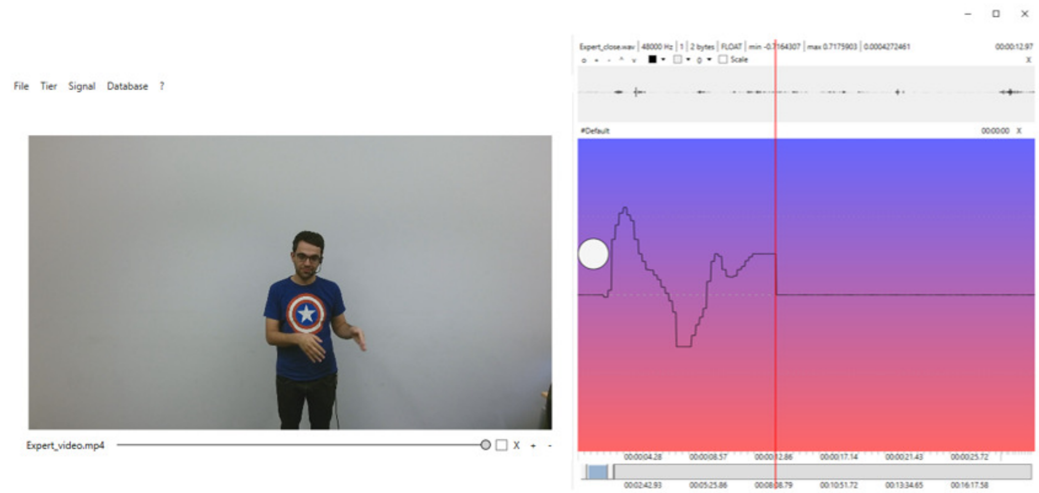

Figure 1: A screen-shot of the interface for continuous annotation in NOVA. For the live annotation mode, a white button is displayed at the left border of the track and only the value at the current playback position (red marker) follows vertical mouse movement (horizontal position of the mouse is ignored). The task is easy and not tiring since it does not require to hold down the right mouse button.

Figure 2: Examples of gestures types: (a) iconic, (b) deictic, (c) metaphoric, (d) beat, (e) object-adaptor, (f) self-adaptor.

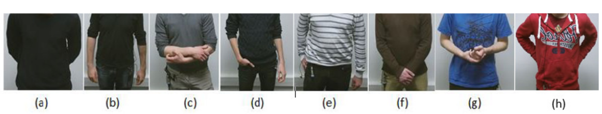

Figure 3: Examples of rest poses: (a) arms_behind, (b) arms_down, (c) arms_crossed, (d) hand_inpocket, (e) hand_onhip, (f) hands_crosseddown, (g) hands_crossedmiddle, (h) hands_onhips.

warmth (or competence) variations: constant, increase and decrease. Each constant was converted in the type of variation immediately preceding it, so that each variation ends when the opposite variation starts. In this way, continuous annotations were converted into binary data. Figure 5 shows an example of this discretization.

The last steps of pre-processing was applied on discrete annotations. First, we merged the annotations coming from the two raters by keeping only the time windows where the two annotators agreed on the type of warmth (or competence) variation expressed by the expert.

Since annotation were sampled at 25 times per second, identical discrete annotations are repeated during the time windows where a non-verbal cue is performed. Thus, we shrank consecutive duplicated samples, yielding the same information, in order to avoid dependency between samples. That is, we kept only samples with at least one different feature or placed in different time windows. The final preprocessed dataset consisted of 1087 samples for warmth and 1069 for competence.

Table 2: The rest poses used in our discrete annotations and their descriptions.

Label	Description
arms_behind	arms are behind the back
arms_down	arms are stretched down along the body
arms_crossed	one arm is put over the other in front of the body, so that each hand is on the opposite elbow
hand_inpocket	one hand is put into the pocket of the trousers, the other one not performing any gesture
hand_onhip	one hand on the corresponding hip, the other one not performing any gesture
hands_crosseddown	arms are laying down, hands are crossed at lower-center level
hands_crossedmiddle	similar to arms_crossed, but only hands are crossed, at center-center level
hands_onhips	two hands on the corresponding hips

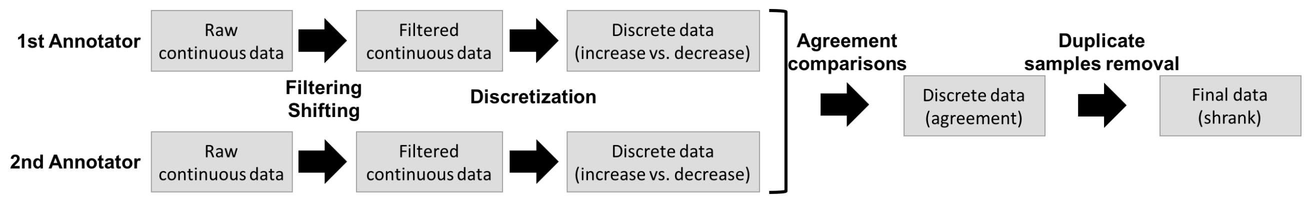

Figure 4: Pipeline of pre-processing of continuous annotations. Data are first processed separately, than last steps are performed only on time windows were the two annotators agreed.

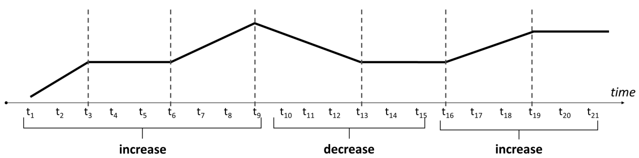

Figure 5: Example of competence variation showing sampled binary discretized levels (increase vs. decrease. When constant, the sample's label for the variation is converted to the same as the one immediately preceding it.

4.2 Analysis and Results

In order to investigate the presence of associations among warmth (or competence) annotations and non-verbal cues, we computed Odds Ratios (ORs) [39]. Odds ratios are an association measure that represents the odds that an outcome will occur given a particular exposure, compared to the odds of the outcome occurring in the absence of that exposure. In our case, they represent the odds that an increase (or a decrease) of warmth (or competence) will occur given a stimulus (type of gesture, or type of rest pose, or smile, or type of head movement), compared to the odds of the decrease (or increase) of the same dimension occurring in the absence of that stimulus. That is:

$$OR = \frac{odds_{increase}}{odds_{decrease}}$$

where

$$odds_c = \frac{p_c}{1 - p_c}$$

and p_c = probability of increase ($1 - p_c$ = probability of decrease) in presence of a non-verbal cue $c \epsilon$ {beat, ideational, adaptor, arms_down, arms_behind, ... , hands_crossed_middle, smile, nod, shake, tilt}.

When $OR = 1$, the presence of the stimulus does not affect odds of increase (no association between the stimulus and warmth -or competence). When $OR > 1$, the presence of the stimulus is associated with higher odds of increase (positive association). When $OR < 1$, the presence of the stimulus is associated with lower odds of increase (negative association with increase, that is, positive association with decrease). A summary of all the computed odds ratios is shown in Tables 3, 4 and 5.

4.2.1 Warmth. **Arms Rest Poses.** In general, the presence of rest poses is associated with decrease of warmth (OR of warmth increase for no_restpose vs all other rest poses = 2.22, p < 0.01). Indeed, when analyzing each rest pose separately, a negative association with warmth is found, in a decrease order of magnitude, for arms_crossed, arms_behind and arms_down.

ORs for hands_on_hips and hands_crossed_down tend towards a positive association with warmth but they do not reach statistical significance.

Type of Gestures. As gestures and rest poses are mutually exclusive, and as the presence of rest poses is associated with decrease of warmth, an increase of warmth is more likely to be elicited in presence of gestures than in their absence. When analyzing each gesture category separately, a high positive association with warmth is found for ideationals and, with smaller magnitude, for beats. No relevant association is found for adaptors.

Head Movements. All head movements show a tendency to be positively associated with warmth, but none of them reaches statistical significance.

Smiling. The highest association with warmth variation concerns smiling: presence of smiling is around 9.7 times more likely to elicit warmth increase compared to absence of smiling (p < 0.0001).

Interaction of smiling and types of gestures. Interesting results emerge from the analysis of gestures performed with a smile or without it. In particular, for ideationals, beats and adaptors, warmth increase is more likely to be elicited when those gestures were made with a smile, compared to without smiling (all ORs > 2.4). The largest effect of smiling is for association between adaptors and warmth: when expert made an adaptor with a smile, raters always annotated warmth increase, while this occurred only in 50% of cases when adaptors were performed without smiling.

Interaction of smiling and rest poses. Smiling positively affects the association of hands_crossed_middle and arms_down (ORs > 10). A warmth increase is more likely to be elicited by these rest poses when they are performed with smile, compared to those without smiling.

Interaction of smiling and heads movements. Tilts, nods and shakes are mostly exhibited without smiling. The fewer cases when they are made with a smile are all associated with warmth increase, while in the other cases only around the 50% are positively associated with warmth.

4.2.2 Competence. **Arms Rest Poses.** In general, the presence of rest poses is associated with decrease of competence (OR of competence decrease for all rest poses vs no rest poses= 1.6, p < 0.001). Specifically, a negative association with competence was found, in a decrease order of magnitude, for arms_crossed and hand_in_pocket.

Table 3: Odds Ratios for arm rest poses, with the correspondent p-value. (n.s. stands for $p > 0.05$, * for $p \leq 0.05$, ** for $p \leq 0.01$, * for $p \leq 0.001$ **** for $p \leq 0.0001$.)**

	no_rest_poses	**arms_down**	**arms_behind**
Warmth	2.2 ****	0.60 **	0.18 ****
Competence	1.6 ****	0.80 n.s.	0.83 n.s.
	arms_crossed	**hands_crossed_down**	**hands_crossed_middle**
Warmth	0.08 **	1.36 n.s.	1.00 n.s.
Competence	0.27 ***	1.46 n.s.	1.00 n.s.
	hands_on_hips	**hand_on_hip**	**hand_in_pocket**
Warmth	3.6 n.s.	0.60 n.s.	1.23 n.s.
Competence	1.5 n.s.	0.90 n.s.	0.4 **

Table 4: Odds Ratios for types of gestures, with the correspondent p-value. (n.s. stands for $p > 0.05$, * for $p \leq 0.05$, ** for $p \leq 0.01$, * for $p \leq 0.001$ **** for $p \leq 0.0001$.)**

	beat	**ideational**	**adaptor**
W	1.4 *	3.09 ***	0.84 n.s.
C	1.6 **	1.3 n.s.	0.88 n.s.

Table 5: Odds Ratios for types of head movements, with the correspondent p-value. (n.s. stands for $p > 0.05$, * for $p \leq 0.05$, ** for $p \leq 0.01$, * for $p \leq 0.001$ **** for $p \leq 0.0001$.)**

	nod	**shake**	**tilt**
W	1.34 n.s.	1.44 n.s.	1.74 n.s.
C	1.5 n.s.	0.94 n.s.	1.08 n.s.

Type of Gestures. As gestures and rest poses are mutually exclusive, and as the presence of rest poses is associated with decrease of competence, an increase of competence is more likely to be elicited in presence of gestures than in their absence. When analyzing each gesture category separately, a moderate positive association with competence is found for beats, and a moderate but no statistically significant association for ideationals. No relevant association is found for adaptors.

Head Movements. Nods show a tendency to be positively associated with competence, but none of head movements reaches statistical significance.

Smiling. Smiling is negatively associated with competence: presence of smiles is 1.6 times more likely to elicit competence decrease compared to absence of smiling ($p < 0.0001$).

Interaction of smiling and types of gestures. Smiling positively affects the association between adaptors and competence: making an adaptor with a smile is 2.21 times more likely to elicit competence increase compared to making an adaptor without smiling. Regarding other gestures, smiling has a moderate negative effect on their association with competence (OR for ideationals = 0.48, OR for beats = 0.37).

Interaction of smiling and rest poses. Smiling positively affects the association of arms_crossed with competence: competence increase is 1.6 times more likely to be elicited by this rest pose when it is performed with smile, than without smiling. In general, for the majority of the other rest poses, smiling has a negative effect on their association with competence.

Interaction of smiling and heads movements. No effects are found for smiling on association between head movements and competence.

4.2.3 Relations between Warmth and Competence. When looking at the direction of the association of each non-verbal cue with the two dimensions of social cognition, we note that for the majority of them an halo effect occurs, while an interesting compensation effect occurs for smiling. Results support the primacy effect of warmth over competence (see Section 2.2). In most of the cases, the magnitude of association (either positive or negative) of each non-verbal cue and warmth is higher than those of the same non-verbal cue and competence. The best evidence for a primacy effect of warmth concerns smiling. The magnitude of association is amplified for warmth (9.67, very high) compared to competence (0.64, moderate in the opposite direction). This is in line with literature [24] where magnitude of compensation effect was found to be higher for warmth compared to competence.

5 DISCUSSION AND FUTURE WORK

In this paper we presented an analysis from a corpus of dyadic expert-novice knowledge sharing interactions. We computed the association between discrete annotations of non-verbal behaviours (type of gestures, arms rest poses, head movements, smiling) with continuous annotations of perceived expert's warmth and competence (converted in two discrete levels describing increases and decreases).

Results show the important role of smiling behaviour. Smile is associated with judgments of warmth increase and competence decrease. This is in line with previous results [9, 17], and suggests evidence of a compensation effect between the two fundamental dimensions of social cognition. Smiling also highly impacts the association of specific types of gestures and rest poses with warmth and competence judgments. For example, when experts were having their arms crossed, competence judgments decreased, but the direction of this association is reversed when the same rest pose co-occurred with a smile. We observed a similar effect between arms crossed and warmth.

The relationship between adaptors (gestures) performed while smiling with increasing competence judgments seems to be in contrast with earlier results. Self-adaptors have been often associated as displays of stress and anxiety [19], that in turn result in a low level of perceived competence. However, our result could be explained by the fact that genuine smiles soften the relationship of the self-adaptors with stress and make more prominent, for the observer, competence perception. Rest poses contributed to decrease in judgments for both dimensions. This is surprising given that in previous work there wasn't any finding linking those behaviours to W&C first impressions.

For the majority of the observed non-verbal cues (except for smiling) we have found evidence in support of the halo effect. More specifically, W&C levels go towards the same direction. Results also support the primacy of warmth over competence in terms of magnitude of effect.

As for head movements, we have found some promising trends (between nods and competence's level and between tilts and warmth) but without reaching statistical significance. We believe that significance can be obtained by annotating more videos, therefore with a larger amount of data.

The videos that we analyzed highlight several interesting and at the same time challenging aspects that are worth further investigation. For example, we observed many differences between interaction styles depending on the interaction participants. Personality seems to play an important role in the expression and perception of W&C, as it does for interpersonal attitudes such as friendliness and dominance [5].

We are aware of some limitations of this work that are related to our methodology. First, continuous annotations of perceived competence level are subjective, thus prone to biases caused by, for instance, annotator's tiredness or social desirability. Secondly, in order to obtain highly reliable annotations with agreement we needed to drastically reduce the amount of data for our analysis, which might have prevented us to find more statistical relationships. A possible solution to overcome the agreement issue is to rely on automatic annotation tools, this also helps adding more discrete annotations, such as gaze and facial expressions. Thirdly, as described in Section 3, verbal behaviour was not considered. In a further step, voice prosodic features as well as speech's content should be collected for looking at the role of both verbal and non-verbal cues in the impression formation of W&C.

The work described in this paper is a first step towards a broader research topic that will include a computational model to manage agent's impressions by generating adaptive nonverbal behavior during the interaction with the user.

In the near future, our goal is first to validate whether user's first impressions of an agent's competence and warmth follow our findings. Then we aim at investigating whether social cognition works in the same way in human-human and human-agent interaction. Moreover, we want to investigate the emotions that expression of warmth and competence through non-verbal cues elicit on users. This can be a valuable information when it comes to detect a user's emotional reaction through non-intrusive objective physiological signals as in our IMPRESSIONS project.

ACKNOWLEDGMENTS

This work is being supported by ANR IMPRESSSIONS project number ANR-15-CE23-0023.

REFERENCES

[1] Andrea E. Abele and Bogdan Wojciszke. 2013. The Big Two in social judgment and behavior. *Social Psychology* 44, 2 (2013), 61–62.

[2] Nalini Ambady and John Joseph Skowronski. 2008. *First impressions.* Guilford Press.

[3] Norman H. Anderson. 1962. Application of an additive model to impression formation. *Science* 138, 3542 (1962), 817–818.

[4] Solomon E. Asch. 1946. Forming impressions of personality. *The Journal of Abnormal and Social Psychology* 41, 3 (1946), 258.

[5] Marieke Astrid, Nicole C. Krämer, and Jonathan Gratch. 2010. How our personality shapes our interactions with virtual characters-implications for research and development. In *International Conference on Intelligent Virtual Agents.* Springer, 208–221.

[6] David Bakan. 1966. The duality of human existence: An essay on psychology and religion. (1966).

[7] Robert F Bales. 1950. A set of categories for the analysis of small group interaction. *American Sociological Review* 15, 2 (1950), 257–263.

[8] Tobias Baur, Gregor Mehlmann, Ionut Damian, Florian Lingenfelser, Johannes Wagner, Birgit Lugrin, Elisabeth André, and Patrick Gebhard. 2015. Context-Aware Automated Analysis and Annotation of Social Human–Agent Interactions. *ACM Transactions on Interactive Intelligent Systems (TiiS)* 5, 2 (2015), 11.

[9] Marjorie A Bayes. 1972. Behavioral cues of interpersonal warmth. *Journal of Consulting and clinical Psychology* 39, 2 (1972), 333.

[10] Kirsten Bergmann, Friederike Eyssel, and Stefan Kopp. 2012. A second chance to make a first impression? How appearance and nonverbal behavior affect perceived warmth and competence of virtual agents over time. In *International Conference on Intelligent Virtual Agents.* Springer, 126–138.

[11] Marino Bonaiuto, Augusto Gnisci, and Fridanna Maricchiolo. 2002. Proposta e verifica empirica di una tassonomia dei gesti delle mani nell'interazione di piccolo gruppo. *Giornale italiano di psicologia* 29, 4 (2002), 777–808.

[12] Angelo Cafaro, Hannes Högni Vilhjálmsson, and Timothy Bickmore. 2016. First Impressions in Human–Agent Virtual Encounters. *ACM Transactions on Computer-Human Interaction (TOCHI)* 23, 4 (2016), 24.

[13] Justine Cassell. 2000. *Embodied conversational agents.* MIT press.

[14] Mathieu Chollet, Magalie Ochs, and Catherine Pelachaud. 2014. From non-verbal signals sequence mining to bayesian networks for interpersonal attitudes expression. In *International Conference on Intelligent Virtual Agents.* Springer, 120–133.

[15] R Cowie and G McKeown. 2010. Statistical analysis of data from initial labelled database and recommendations for an economical coding scheme. *SEMAINE Report D6b* (2010).

[16] Amy JC Cuddy, Susan T Fiske, and Peter Glick. 2008. Warmth and competence as universal dimensions of social perception: The stereotype content model and the BIAS map. *Advances in experimental social psychology* 40 (2008), 61–149.

[17] Amy JC Cuddy, Peter Glick, and Anna Beninger. 2011. The dynamics of warmth and competence judgments, and their outcomes in organizations. *Research in Organizational Behavior* 31 (2011), 73–98.

[18] d B Duchenne. 1990. The mechanism of human facial expression or an electro-physiological analysis of the expression of the emotions (A. Cuthbertson, Trans.). *New York: Cam-bridge University Press.(Original work pub-lished 1862)* (1990).

[19] Paul Ekman and Wallace V Friesen. 1974. Nonverbal behavior and psychopathology. *The psychology of depression: Contemporary theory and research* (1974), 3–31.

[20] Susan T Fiske, Amy JC Cuddy, and Peter Glick. 2007. Universal dimensions of social cognition: Warmth and competence. *Trends in cognitive sciences* 11, 2 (2007), 77–83.

[21] Susan T Fiske, Amy JC Cuddy, Peter Glick, and Jun Xu. 2002. A model of (often mixed) stereotype content: competence and warmth respectively follow from perceived status and competition. *Journal of personality and social psychology* 82, 6 (2002), 878.

[22] Susan T Fiske and Steven L Neuberg. 1990. A continuum of impression formation, from category-based to individuating processes: Influences of information and motivation on attention and interpretation. *Advances in experimental social psychology* 23 (1990), 1–74.

[23] Erving Goffman and others. 1978. *The presentation of self in everyday life.* Harmondsworth.

[24] Charles M Judd, Laurie James-Hawkins, Vincent Yzerbyt, and Yoshihisa Kashima. 2005. Fundamental dimensions of social judgment: understanding the relations

between judgments of competence and warmth. *Journal of personality and social psychology* 89, 6 (2005), 899.

[25] Ashish Kapoor and Rosalind W Picard. 2001. A real-time head nod and shake detector. In *Proceedings of the 2001 workshop on Perceptive user interfaces*. ACM, 1–5.

[26] Nicolas Kervyn, Hilary B Bergsieker, Fiona Grignard, and Vincent Y Yzerbyt. 2016. An advantage of appearing mean or lazy: Amplified impressions of competence or warmth after mixed descriptions. *Journal of Experimental Social Psychology* 62 (2016), 17–23.

[27] Fridanna Maricchiolo, Augusto Gnisci, Marino Bonaiuto, and Gianluca Ficca. 2009. Effects of different types of hand gestures in persuasive speech on receivers' evaluations. *Language and Cognitive Processes* 24, 2 (2009), 239–266.

[28] Soroosh Mariooryad and Carlos Busso. 2015. Correcting time-continuous emotional labels by modeling the reaction lag of evaluators. *IEEE Transactions on Affective Computing* 6, 2 (2015), 97–108.

[29] David McNeill. 1992. *Hand and mind: What gestures reveal about thought*. University of Chicago press.

[30] Margaret McRorie, Ian Sneddon, Etienne de Sevin, Elisabetta Bevacqua, and Catherine Pelachaud. 2009. A model of personality and emotional traits. In *International Workshop on Intelligent Virtual Agents*. Springer, 27–33.

[31] Albert Mehrabian. 1969. Significance of posture and position in the communication of attitude and status relationships. *Psychological Bulletin* 71, 5 (1969), 359.

[32] Angeliki Metallinou and Shrikanth Narayanan. 2013. Annotation and processing of continuous emotional attributes: Challenges and opportunities. In *Automatic Face and Gesture Recognition (FG), 2013 10th IEEE International Conference and Workshops on*. IEEE, 1–8.

[33] Truong-Huy D Nguyen, Elin Carstensdottir, Nhi Ngo, Magy Seif El-Nasr, Matt Gray, Derek Isaacowitz, and David Desteno. 2015. Modeling Warmth and Competence in Virtual Characters. In *International Conference on Intelligent Virtual Agents*. Springer, 167–180.

[34] Radosław Niewiadomski, Virginie Demeure, and Catherine Pelachaud. 2010. Warmth, competence, believability and virtual agents. In *International Conference on Intelligent Virtual Agents*. Springer, 272–285.

[35] Guido Peeters. 2002. From good and bad to can and must: Subjective necessity of acts associated with positively and negatively valued stimuli. *European Journal of Social Psychology* 32, 1 (2002), 125–136.

[36] Catherine Pelachaud. 2009. Modelling multimodal expression of emotion in a virtual agent. *Philosophical Transactions of the Royal Society of London B: Biological Sciences* 364, 1535 (2009), 3539–3548.

[37] Brian Ravenet, Magalie Ochs, and Catherine Pelachaud. 2013. From a user-created corpus of virtual agentâĂŹs non-verbal behavior to a computational model of interpersonal attitudes. In *International Workshop on Intelligent Virtual Agents*. Springer, 263–274.

[38] Seymour Rosenberg, Carnot Nelson, and PS Vivekananthan. 1968. A multidimensional approach to the structure of personality impressions. *Journal of personality and social psychology* 9, 4 (1968), 283.

[39] Nova Scotia. 2010. Explaining odds ratios. *J Can Acad Child Adolesc Psychiatry* 19 (2010), 227.

[40] Janine Willis and Alexander Todorov. 2006. First impressions making up your mind after a 100-ms exposure to a face. *Psychological science* 17, 7 (2006), 592–598.

[41] Bogdan Wojciszke and Andrea E Abele. 2008. The primacy of communion over agency and its reversals in evaluations. *European Journal of Social Psychology* 38, 7 (2008), 1139–1147.

[42] Bogdan Wojciszke, Roza Bazinska, and Marcin Jaworski. 1998. On the dominance of moral categories in impression formation. *Personality and Social Psychology Bulletin* 24, 12 (1998), 1251–1263.

[43] Yi-Hsuan Yang and Homer H Chen. 2011. Ranking-based emotion recognition for music organization and retrieval. *IEEE Transactions on Audio, Speech, and Language Processing* 19, 4 (2011), 762–774.

[44] Vincent Y Yzerbyt, Nicolas Kervyn, and Charles M Judd. 2008. Compensation versus halo: The unique relations between the fundamental dimensions of social judgment. *Personality and Social Psychology Bulletin* 34, 8 (2008), 1110–1123.