A Systematic Review on the Socio-affective Perception of IVAs' Multi-modal behaviour

Elodie Etienne
QuantOM, HEC Management School
University of Liège
Liège, Belgium
elodie.etienne@uliege.be

Marion Ristorcelli
LIS and CNRS, Aix-Marseille
University
Marseille, France
marion.ristorcelli@lis-lab.fr

Sarah Saufnay
QuantOM, HEC Management School
University of Liège
Liège, Belgium
Sarah.Saufnay@uliege.be

Aurélien Quilez
LIS and CNRS, Aix-Marseille
University
Marseille, France
afasolo@u-bordeaux.fr

Rémy Casanova
ISM and CNRS, Aix-Marseille
University
Marseille, France
remy.casanova@univ-amu.fr

Michaël Schyns
QuantOM, HEC Management School
University of Liège
Liège, Belgium
M.Schyns@uliege.be

Magalie Ochs
LIS and CNRS, Aix-Marseille
University
Marseille, France
magalie.ochs@lis-lab.fr

ABSTRACT

The multimodal behaviour of IVAs may convey different socio-affective dimensions, such as emotions, personality, or social capabilities. Several research works show that factors may impact the perception of the IVA's behaviour. This paper proposes a systematic review, based on the PRISMA method, to investigate how the multimodal behaviour of IVAs is perceived with respect to socio-affective dimensions. To compare the results of different research works, a socio-emotional framework is proposed, considering the dimensions commonly employed in the studies. The conducted analysis of a wide array of studies ensures a comprehensive and transparent review, providing guidelines on the design of socio-affective IVAs.

CCS CONCEPTS

• **Human-centered computing** → **User studies**.

KEYWORDS

Virtual agent, multimodal behaviour, socio-affective dimensions, perception, systematic review

1 INTRODUCTION

Intelligent Virtual Agents (IVAs) are able to display a wide range of multimodal behaviours to interact naturally with the users in a virtual environment, whether in 2D or 3D. Depending on the role endowed by the IVAs (e.g., virtual guide [4], virtual recruiter [5], or virtual patient [37]), they may be required to convey different socio-affective states such as a dominant attitude, positive emotions, or a behaviour that inspires confidence. Several research works show that both the verbal and the non-verbal behaviour of the IVA strongly impact the users' perception of the IVA's socio-affective state. For instance, IVA may display different facial expressions such as smiles, frowns and raised eyebrows to convey a specific emotional state during the interaction [4–6, 9, 14, 18, 22, 23, 26, 28, 32, 38, 40, 42, 45, 47, 49]. The IVA can also direct or avert its gaze from the interlocutor to display different levels of engagement [5, 6, 9–11, 14, 18, 26, 32, 35, 45, 47]. The other modalities of non-verbal behaviours are the head movements (shaking, nodding, or tilting) [6, 14], the orientations of the head (downward, upward, inclined or straight) [5, 53], the torso positions (leaning forwards, backwards or straight) [32], the arms positions (e.g., crossed arms, arms behind the head, hands on the hips), and the body position (e.g., standing or sitting on a chair) [4, 35, 43, 47]. Furthermore, the breadth of movements (from small to large radius), their intensity (light, moderate, or forceful) [4], and the directionality of gestures (pointing with a finger or an open palm) further enrich this non-verbal vocabulary to convey socio-emotional state [5, 53]. Hand gestures also play a crucial role, ranging from central positioning to peripheral actions, and alternating between behaviours like fiddling with hands, and shrugging [4, 35], or hands clasped on a table.

Regarding the verbal modalities, there exists a variety of linguistic and paralinguistic signals. Linguistic signals pertain to the elements of language and communication that can be explicitly observed and analysed. These include the choice of words, the structure of sentences, grammar, and syntax. Some specific examples are the use of subjective pronouns, verbs, nouns, the level of formality of the language, the use of self-references, the variation of vocabulary, the length of the sentences, the use of positive or negative contents [3–5], and the use of commanding or suggesting sentences [53]. Paralinguistic signals, although closely related to linguistic ones, extend beyond the basic grammar and syntax to include aspects of meaning that are not directly encoded in the linguistic elements. These include acoustic signals such as the pitch and the speech rate [16], the behaviour impacting the flow of the conversation, as for example the overlapping of speech when the participant interrupts the IVA [17]. Another important aspect is the alignment or coordination of the behaviours of the two interlocutors, considering, for instance, the sequentiality and temporality of signals [10, 38] to develop IVAs capable to display different attitude variations or to adapt to the user's perception, as proposed in [4].

Furthermore, several factors can have an impact on the socio-affective perception of the IVA's behaviour, some related to the IVA (e.g., its appearance [7, 11, 23, 32]), others related to the user (e.g., age) [23].

Although there are many studies on this subject, it remains difficult to have a clear vision of the different socio-affective behaviours that IVAs can convey and how they can express them. *The objective of the paper is precisely to provide a systematic review, based on the PRISMA method [33, 50], of the research works investigating the users' perception of the socio-affective dimensions of IVAs conveyed through their multimodal behaviours.*

Researchers have explored a wide range of socio-affective dimensions that IVAs may convey through their behaviour. In order to compare the results of different studies, the socio-affective dimensions have been gathered into five categories: emotions, personality, trustworthiness, social capabilities, and believability. The *emotions* category groups the perceptive studies on the users' perception of the emotions or the moods expressed by the IVAs. The *personality* category includes the studies on the users' perception of the IVAs' personality traits such as friendliness, dominance, and extroversion. The *trustworthiness* category gathers perceptive studies on the impact of IVAs' behaviour on the perception of trust, including the perception of competence, intelligence and cooperativeness. The *social capabilities* category covers studies on the perception of a virtual relationship for the creation of a social connection with the user or between IVAs [19]. The last category is dedicated to the perceived *believability* of the IVAs. Of course, the proposed socio-emotional framework is subject to scrutiny. The categories are closely linked to each other, e.g., emotions are influenced by personality [12, 34], some personality traits are sometimes considered as emotional dimensions [45]. However, this socio-emotional framework - constructed by grouping the socio-affective dimensions mentioned in the papers according to their proximity in terms of definition and use - enables a comprehensive and transparent review of the existing works presented in this paper.

This systematic review is guided by a central research question: *How are the emotions, personality, trustworthiness, social capabilities,*

and believability of IVAs perceived by users through their multimodal behaviour? Indeed, the objective is to identify more precisely the signals that IVAs can use to convey these socio-affective dimensions, but also the importance of each modality depending on the considered socio-affective dimension, and the effects of the combinations of signals on perception. Moreover, the aim is to highlight the different factors - related to the IVA, to the user, and to the interactive device - that may influence the perception of the user.

This paper is structured as follows. The next section (Section 2) outlines the approach, based on the PRISMA method [33, 50], to perform the systematic review. Subsequently, Section 3 describes the results of the papers included in the systematic review on users' perception of the five socio-affective categories introduced above through the IVAs' multimodal behaviour. Finally, Section 4 delves into the discussion of the variability in the perception of multimodal behaviour influenced by various factors.

2 METHODOLOGY

The PRISMA method [33, 50] is used to guide the systematic review process. This method ensures transparency and rigour by systematically identifying, assessing, and evaluating relevant literature. The focus is on the perceptual review of how the verbal and non-verbal behaviours of IVAs influence the perception of emotions, personality, trustworthiness, social capabilities, and believability. The criteria applied in the selection process include studies on the perception of the socio-affective dimensions of at least one IVA depicted in virtual environments based on monitor or VR.

Thus, the literature search conducted in the Scopus database[1] uses the following request:

```
(virtual AND (audience* OR avatar* OR agent* OR
listener* OR character*)) AND (perception OR perceive*)
AND (behaviour* OR behavior* OR "body language" OR
arousal OR valence OR stance OR attitude*) AND (nonverbal
OR non-verbal OR verbal) NOT patholog* NOT autism.
```

This search aims to identify papers exploring the perception and behaviours of IVA, excluding those related to pathology and autism.

The search process retrieves a total of 162 papers in Scopus. Exclusion criteria are defined to refine the search results, ensuring relevance and focus on the research question. The criteria encompass the exclusion of papers with accessibility issues, duplicates entries, non full papers or papers without any DOI, papers that are not a perceptive study. Furthermore, as the focus is on the perception of IVA using multimodal behaviour, papers in which the IVA does not have a face or a humanoid body are removed. The same happened for papers that present neither verbal nor non-verbal behaviour for the IVA. With the focus on virtual environments using monitor or VR, papers not using monitor or VR are excluded . Finally, the emphasis is placed on adult Occidental participants. Consequently, all studies involving participants with a mean age below 18 years old or non-Occidental participants are also excluded[2].

The first step of the PRISMA method consists of selecting articles by reading their title and abstract. From the 162 articles identified using the query presented above, the initial screening leads to the

---

[1]In fact, the search was also performed using Web of Science, resulting in 98 papers. However, since these papers were already included in Scopus, they were not utilised.
[2]For studies conducted in Occidental universities or research groups, the absence of such participants cannot be definitively confirmed when full details were not given.

exclusion of 73 papers. Subsequently, a more thorough examination of the remaining papers is conducted, involving reading through each in its entirety, which constitutes step 2 of the screening process. After this comprehensive review, an additional 57 papers are excluded. See Table 1 for the more details.

This methodical filtration underscores the rigorous selection and exclusion criteria inherent in the PRISMA approach, ensuring that the review focuses on the studies most relevant to the research question. Table 2 proposes a summary of the dimensions, sub-dimensions, and signals involved as dependent variables for each selected paper. In the next section, the results of the paper for each socio-affective category are described in more detail.

3 RESULTS

3.1 Emotions

In this section, the research works investigating the perception of IVA's emotions through their multimodal behaviour are reported. Initially, the literature reveals that emotional dimensions are explored through different affective concepts such as mood and emotions [1]. To compare the research results, the analysis of the research works on the emotional perception of IVA behaviour is proposed in light of the *valence* and *arousal* dimensions.

These dimensions are expressed and perceived in various ways, for example, through posture and facial expressions [24]. According to the authors, the SAM questionnaire, which studies the dimensions of valence, arousal, and dominance, is a good tool to assess the perception of these dimensions. However, dominance, which could also be considered as a personality trait, is discussed in Section 3.2.

According to [6], the notion of valence refers to the IVA's opinion, *i.e.*, the positive or negative feelings it has towards the user. As shown in several research works [6, 9, 14, 18, 28, 38], the emotional state, and more specifically valence, is conveyed by facial expressions to display basic emotions such as anger, happiness, and sadness [28], but also more subtle emotional states such as stress [9], amusement, or politeness [38]. Even if facial expressions are a strong cue of emotional state, the way they are displayed in IVAs can vary considerably from one study to another, leading to misinterpretation depending on the considered representation, the intensity of the expression, and the combination of modalities [14, 28]. Although it seems that the frown conveys a negative valence [9, 13], the way it is represented in the IVA could vary, particularly in terms of intensity and of the parameters used to display facial expression. For example, in the study of [14], a smiling facial expression is perceived as neutral, while it is generally a positive valence signal, indicating a potential confusion between a fake smile and a genuine smile [14, 38]. Valence is mainly conveyed by the combination of facial expressions and head movements. Indeed, researchers consistently associate smiling and nodding with positive valence, and frowning and shaking the head with negative valence [6, 9, 14, 18]. However, some signals may be predominant in the perception of valence. Indeed, as shown in [14], head shake is the most negative modality identified and is always judged negatively, regardless of the other modality it is associated with. Regarding the other signals, the literature is not consensual. For example, posture, such as crossed arms, may convey a negative valence as shown in [6, 14], while [18] does not find no impact of posture on the

assessment of valence. To express the valence, some modalities may be more important than others. For instance, as highlighted in [6], to assess the IVA's valence, the users generally first consider the head movements, then the posture, the gaze direction, and finally the facial expressions, as this signal is more subtle in VR. Additionally, research on the perception of behaviour on the valence highlights the importance of combining non-verbal behavioral modalities. For example, nodding is mainly a sign of positive valence but is sometimes considered as neutral based on the associated signals [14]. In the same way, an IVA with the head tilted, its elbow on the table and its torso leaning forward appears to be positive, while separately, these signals convey a neutral or negative valence [14]. The perception of the non-verbal modality may also vary according to the associated verbal behaviour, as highlighted in [38]. In addition, vocal non-verbal immediacy, corresponding to the pitch level and speech rate, may also have an impact on the assessment of participants' affect towards the content or the virtual model. It also impacts the likelihood of following the same virtual instructor again for other similar videos in the future [16]. Indeed, stronger vocal immediacy, characterised by an average pitch of 260 Hz and a speech rate of 133 words per minute (wpm), enhance affective learning compared with a virtual model that uses weaker vocal immediacy, with an average pitch of 115 Hz and a speech rate of 119 wpm [16]. Some factors may influence the perception of valence, such as the appearance of the IVA itself. Although few studies examine this factor, some research shows that a female IVA is perceived as more positive than a male IVA when smiling [38]. Another study also shows that for a given facial expression, such as raising eyebrows, a virtual female IVA appears to be more stressed if it uses a frontal body with direct gaze than if it used an averted body with an averted gaze, whereas no difference is found between these behaviours for the male IVA [9].

The dimension of arousal describes the excitement of the event for a person and ranges from low to high alertness [51]. It takes the form of an IVA interested or attentive to what people are saying and is characterised by two types of non-verbal behaviour, namely proximity and body movements [51]. Consistently, all the authors who study this dimension find a relationship between the evaluation of arousal and the direction of gaze, the frequency of movements, and the proximity of posture [6, 14, 18]. A high level of arousal or engagement is associated with an IVA looking at the speaker, with a closer posture, *i.e.* a torso leaning forward and frequent head movements and facial expressions. On the contrary, an IVA that looks away with a more distant and relaxed posture is perceived as disconnected [6, 18]. It is also interesting that eyebrow raising is judged to be neutral in terms of arousal and that head shaking is always associated with high arousal, regardless of other signals [14]. Another interesting point is the relationship between perception of valence and arousal. The combination of the valence-arousal pair can convey a specific attitude, and, as indicated by [18], users are able to perceive different social attitudes based on these dimensions (indifferent, critical, and enthusiastic). However, several authors report that users do not distinguish between different levels of valence (positive and negative) for low arousal [6, 14, 18].

3.2 Personality

In this section, three personality traits are considered: the *dominance*, the *extroversion*, and the *friendliness*. The dominance is part of the PAD (Pleasure-Arousal-Dominance) model, described by [44]. The dominance corresponds to a scale ranging from the absence of control or impact on the event to the feeling of influence or control on the situation [51]. Extroversion and friendliness are two components of the "Big Five" personality traits model [21] that identifies Openness, Conscientiousness, Extroversion, Agreeableness, and Neuroticism. As explained in [11], agreeableness is an indicator of friendliness. Warmth may also refer to friendliness [27]. Similarly, the concept of politeness involves the dimension of friendliness, as explained in [52, 53]. For the sake of clarity, only the term *friendliness* is used in this section. According to [1], extroversion corresponds to the sociability of IVA perceived by users.

The choice of only keeping dominance, extroversion and friendliness to express personality is explained by the interrelationships between them. Indeed, [3] shows that dominance and extroversion are positively correlated, while dominance and friendliness are negatively correlated. In addition, friendliness and dominance are two dimensions that appear in the Interpersonal Circumplex model [27], which further indicates their strong relationship.

To express dominance, an IVA can look at the user in various ways: by looking from below, which is perceived as less dominant, from above, or by aligning its eyes close to those of the participant [3]. Furthermore, as shown by [43, 47], dominant non-verbal behaviours can be displayed using akimbo posture, crossing arms, sagittal head up, gesture with large radius. On the contrary, submissive non-verbal behaviour includes neck-adapter (self-touch), arms open, sagittal head down, gesture with small radius.

Regarding non-verbal signals, the more dominant position is adopted, the more dominant the IVA is perceived [43]. Furthermore, [47] shows that, concerning the perception of dominance, crossing arms and the akimbo position are the most effective gestures, but taking up more space (gesture with a large radius) is not perceived as more dominant than the other gestures. Furthermore, [11] and [10] demonstrate that by using more (or less) dominant cues, an increase (or decrease) in perceived dominance can indeed be implied. As shown by [3], the more linguistic friendly cues are used, the less dominant is the IVA perceived. It is further established by [5] when using verbal modalities of friendliness. Moreover, regarding vocal modalities, [17] shows that the longer the interruption handling time, the more dominant the IVA was perceived.

To express extroversion, IVAs can exhibit positive emotional states and medium arousal levels (see Section 3.1), and high dominance behaviours (closer interaction distances, expansive gestures, rapid gesture execution, sustained eye contact, and prolonged gaze duration). In contrast, introverted agents display negative emotional states, low arousal, more distant interactions, reduced spatial extent in gestures, slower gesture speeds, and averted gazes [7, 45].

As shown by [3], the more dominant gaze cues are used, the less extroverted the IVA is perceived. Generally, it is possible to distinguish extroversion from introversion by voice or facial expression alone [46]. However, if the voice is combined with body movements, it is the most informative signal to judge the extroversion of the virtual agent [46]. More specifically, [45] shows that a virtual agent is considered to be extroverted if it has a positive emotional state, a medium level of arousal (see Section 3.1), and a positive dominance value. On the contrary, an introverted virtual agent has a negative emotional state, a low level of arousal, and a low dominance value. At the behavioural level, extroverted personality translates into a closer distance during the interaction (e.g., the torso leaning forward), a higher value of spatial extent during the execution of gestures, a higher speed of execution gestures with greater eye contact, and a longer duration of gaze for an extroverted virtual agent. On the contrary, an introverted virtual agent is more distant during interaction (leaning back) and adopts a lower value of spatial extent and speed of gesture execution, with an averted gaze [7].

To express friendliness, regarding verbal modalities, an IVA can use fewer synonyms and negations, shorter sentences, more pronouns, verbs, negations, informal language, positive content, and references to the speaker [4].

Regarding non-verbal modalities, the more dominant positions are adopted, the less friendly the IVA is perceived [43]. As expected, IVA friendliness can be perceived through positive facial expression (e.g., smile [38]) but specific gestures, such as commanding the user through arm gestures (using finger pointing) can have quite the reverse effect [52]. However, [5] shows that non-verbal behaviour or verbal modalities of friendliness alone are not sufficient to render friendliness, suggesting that verbal modalities must be added to ensure the right perception of it. Furthermore, [52] shows that, to maintain a certain degree of friendliness in the IVA, it is preferable to use suggestion rather than command. In addition, one may prefer stronger vocal immediacy (*i.e.* higher pitch and faster speech rate) to depict a more friendly IVA [3]. Lastly, regarding vocal modalities, [17] shows that the longer the interruption handling time, the less friendly the IVA is perceived. Moreover, [10] and [11] show that using more (respectively less) dominant cues and/or less (respectively more) friendly cues implies an increase (or decrease) of perceived friendliness. Similarly, [4] shows that the model that uses adaptive algorithm (Reinforcement Learning) to adapt to user impressions can indeed increase the perceived degree of friendliness compared to when IVA does not adapt its behaviour to user reactions.

3.3 Trustworthiness

Trustworthiness is a complex dimension, closely related to underlying concepts, such as the performance of the IVA and its inclination to collaborate [20]. In this literature review, the perception of several sub-dimensions of trust is investigated, such as the IVA's *competence*, *intelligence*, *autonomy* and *helpfulness*, reflecting its performance, and its cooperativeness and persuasiveness, displaying a certain predisposition for collaboration with the user.

In the literature, research highlights the impact of IVA's emotional non-verbal behaviour on its attributed *trust* level. In [23], the IVA displays an emotional non-verbal behaviour, either positive (*i.e.* smile, head nod, head nod plus smile) or negative (*i.e.* sad face, head down, dropping the arms plus sad face), while respectively announcing good or bad news. Compared to an IVA that remains neutral in its behaviour, its perceived trustworthiness increases under emotional conditions. This relationship is further studied in [47], more specifically by investigating the impact of emotional behaviours on perceived cooperativeness. The hypothesis

assumed by the authors, justifying the adoption of such behaviours to depict cooperativeness, states that this sub-dimension is closely related to IVAs' expressiveness. This hypothesis is confirmed, therefore strengthening the already established link between an IVA's emotions and trustworthiness. The results of the considered study include the preeminence of behaviours combining expressive gestures and mimic to provide the IVA with a cooperative attitude. The impact of lateral head tilts, whatever the side, is positive as well. Naturally, opposite results are found for neutral and non-expressive behaviours, which are associated to lower cooperativeness perception levels, however not as low as averting gazes.

The personality of the IVA, either depicted by its friendly or dominant behaviour towards the user, is also recognised as primordial in shaping perceived competence, trustworthiness, and cooperativeness [4, 10, 35, 43, 47]. A model commonly used in the literature is the *Warmth and Competence* model [15], which links the IVA's friendliness (see 3.2) with its competence. This model is adopted by [4] and [35] to assign specific non-verbal behaviours to IVAs depending on the desired competence level. High competence perception is successfully obtained by synchronising the IVA's gestures with the semantic content of the speech [35]. In contrast, IVAs convey low competency through desynchronised gestures [35]. This manipulation effectively transmits the expected competence signal, regardless of the IVA warmth level, which is characterised by an open (high warmth) or closed gesture (low warmth). However, it is worth mentioning that IVA's warmth, and by extension its friendliness, also has an impact on competence perception, with a positive correlation between these two variables. The accuracy of the warmth and Competence model to design IVAs perceived as such is confirmed in a second study [4].

As previously stated, studies also investigate the interrelation between dominant behaviours and trust sub-dimensions, more specifically considering the impact on competence, intelligence and persuasion [43], as well as on helpfulness [10] and cooperativeness [10, 43, 47]. The effect of dominance is, however, not confirmed for all sub-dimensions. In [43], classical dominant non-verbal signals (*i.e.* akimbo posture, arms crossed, gestures with large radius, sagittal head up) have no effect on competence, autonomy, intelligence, and persuasion ratings. On the contrary, the helpfulness of the IVA appears to be negatively impacted [10]. When it comes to cooperativeness perception, contradictory results can, however, be observed. Despite [43] additionally confirms that no relationship can be established between these two variables, [47] finds a negative effect of dominance on cooperativeness, although its effect remains quite small. The latter study additionally identifies specific dominant signals, particularly perceived as uncooperative, such as keeping the arms crossed, which corresponded to the lowest cooperativeness rating among several dominant behaviours, including keeping the head up or making large radius movements. Conversely, the authors identify IVA' behaviours, associated to submission, interpreted as cooperative signals. More specifically, an IVA adopting small radius gestures or keeping its arms open improves its attributed cooperativeness level.

The persuasiveness of IVAs' multimodal behaviour is also investigated. Two studies firstly compare the effectiveness of several strategies in order to persuade the user to join a group of IVAs, either using monitor [52] or a VR headset [53]. In both conditions,

similar results are obtained. Strategies adopting a direct approach, explicitly formalising the request, are the most persuasive ones. Actually, when the IVA directly commands the user to move to a specific location, while pointing it with its index, the level of associated persuasiveness is the highest among all possible strategies [52]. In comparison, less commanding approaches, such as politely asking or proposing to the user to join the group, are less persuasive but yet still effective [52]. The authors further highlights the primordial importance of clear formulation of the demand to maximise IVAs' force of persuasion. In addition to the influence of the strategy adopted, communication modalities also play a role. Indeed, multimodal modalities, combining verbal (*i.e.* proposition or command) and non-verbal communication (*i.e.* gaze and hand movements), reach the highest impact levels on persuasiveness [53]. Nevertheless, in [39], results indicate that verbal modality only should be preferred to maximise the IVA's force of persuasion, reducing distractions for the user, who can therefore be focused on the command itself. An important difference between this study and previous ones that should be considered is that the non-verbal behaviour in this case is not directly related to the command asked of the user, thereby justifying its perceived uselessness.

Generally speaking, the IVA's verbal behaviour also plays a significant role in trust perception. The importance of a realistic voice is highlighted, with IVAs eliciting higher level of trust when endowed with a human voice rather than a synthetic one [39]. Verbal behaviour also turns out to be important in eliciting competence. An interesting signal used to reflect high competence and proved to be relevant is the formulation of sentences using *"We"* or *"You"* pronouns, rather than sentences formulated in the first person singular [4]. Similarly, IVAs disclosing personal information are not perceived as more competent [48], further reinforcing the previous statement regarding the irrelevance of *"I"* pronouns to elicit competence. Although the words used are important, the number of words used has no influence on competence perception [48].

Valuable insights for the design of trustworthy IVAs are provided, but the design of their behaviours should still be carefully considered. Indeed, specific aspects can considerably impact perception, leading to undesired effects. Such negative effects are observed in [49], with mimicking IVAs, evoking a low level of trust and of helpfulness among users when reproducing their behaviour, being even perceived as creepy. The immediacy of the reaction further reinforces the identified negative effect. Another surprising result from the considered papers. [43] identifies that guiding behaviours reduce the perceived level of competence. Such behaviours are represented with deistic gestures and gaze, meaning that IVAs pointing to specific elements are not seen as competent.

3.4 Social Capabilities

In this section, the research works exploring users' perception of IVAs' social capabilities are presented. The articles considered in the systematic review identify the following main dimensions related to social capabilities: the *mutual understanding* or *comprehension*, the *mutual agreement*, the *intimacy* or *self-disclosure*, the *interpersonal and dyadic stances* and finally, the *social status*. These dimensions are strongly related to the notion of virtual rapport [19].

The *mutual understanding, attention, agreement, interest and pleasantness* are explored in [42] in terms of users' perception of an interaction between two IVAs. In social interaction, intimacy corresponds to "a reciprocal expression of personal or emotional contents, and the perception of positive feelings and comprehension." [41]. In [25], intimacy refers to the perceived self-disclosure of the IVA, which corresponds to the capacity to share personal information with the user to create social connections. The interpersonal stances in [38] correspond to the perceived relationship expressed by the IVA toward its interlocutor, as for instance cold or warmth. Finally, the dyadic stance results from a behaviour alignment or nonalignment between the two agents reflecting for instance an agreement or hostility [38].

The perception of these social capabilities dimensions described above can be significantly influenced by the multimodal behaviour of the IVA. For instance, as shown in [42] and [38], the mutual reinforcement of smiles between two IVAs or between an IVA and a user, enhances the perception of the *mutual understanding* and impacts the perception of the interpersonal and dyadic stances [38].

Concerning *intimacy*, in [40, 41], the authors show that a high level of perceived intimacy leads to a longer interaction with the user. Moreover, the behaviour of the agent may influence its perceived honesty and genuineness: users rated the IVA as more honest and authentic when it displays behaviours associated with intimacy such as emotional facial expressions (e.g., smiling), open-arm gestures, self-directed motions, head nods and tilts, and eye contact. Moreover, [48] investigates the perception of the talkative and self-disclosure dimensions. Results show that an IVA is perceived as more talkative and open, when it uses more words and shares personals information with the user.

A last sub-dimension is the influence of its non-verbal behaviours on its perceived social status. Among the selected papers in this literature review (see Table 2), a unique study, [36] aims to investigate the effects of different gaze behaviours, reflective of either high or low social status, during a job interview scenario with an IVA. The study finds that an IVA perceived as belonging to a higher social status displays longer duration of eye contact and prolongs stares following the user's responses. Furthermore, an IVA raising his head posture and with a body leaned toward the user is also associated with a higher social status, highlighting the significance of non-verbal modalities in shaping perceptions of social hierarchies within virtual interactions.

3.5 Believability

The user's perception may be strongly impacted by the believability of the IVAs influenced by the quality of the animations and of the voice. In this literature review, believability is assessed across several dimensions. In [16], the authors refer to the dimensions of *animacy*, *anthropomorphism* and *liveliness* which correspond to dimensions of the Godspeed questionnaire [2]. The notion of believability corresponds also to the perception of *credibility*, for instance in [5] to explore how credible the IVA is in its role of recruiter. In [26], the believability of the IVA's behaviour is evaluated through the notion of *plausibility* and *naturalness* focusing on both the behaviour and the appearance.

Several articles in our literature review have shown the influence of IVAs' behaviours on the users' perception of believability. First of all, concerning the animations, [39] and [26] emphasise that an animated IVA, regardless of the animation mode, is perceived as more natural and more plausible than a static one, indicating that the movements are a key factor in the perception of believability. Moreover, [39] shows that a more varied and nuanced behaviour leads to a better perception of the realism. [26] also focuses on the naturalness and plausibility of facial animation behaviour, showing that synthesised expressions (*i.e.* facial expressions generated from data such as audio, head movements, and tagged gaze targets to correspond with expected facial expressions in specific situations) are evaluated as more natural and plausible than tracked expressions (*i.e.* facial expressions captured in real-time from facial movements), both for verbal and non-verbal behaviour. In [29], the authors highlight the importance of the appropriateness of the expressed signals showing that inappropriate nods are perceived as less natural than those adhering to established norms.

In [16], the authors investigate the impact of the verbal behaviour and in particular the vocal immediacy on the animacy, anthropomorphism and liveliness. They manipulate the virtual agent's vocal parameters, such as pitch and speech rate. Results indicate that participants exposed to the condition with stronger vocal immediacy (high pitch and fast speech rate) perceive the agent as more anthropomorphic, animated, and likeable compared to those in the weaker vocal immediacy condition.

4 DISCUSSION

In this paper, the interplay of socio-affective perceptions of IVAs by users is delved into, specifically focusing on how emotions, personality, trustworthiness, social capabilities, and believability are discerned through the IVAs' multimodal behaviour. To achieve this goal, the PRISMA method is employed, ensuring a systematic and transparent review process that underpins the analysis.

This exploration reveals a rich tapestry of dimensions used across studies to discuss socio-affective perception. However, a detailed analysis of the dimensions described in the papers shows that the diverse terminology are used to describe similar or overlapping concepts. This diversity poses challenges but also offers an opportunity for synthesis. For example, as explained in Section 3.2, several terms can be used to describe friendliness (e.g., agreeableness, politeness, warmth). In this paper, a novel comparison of papers included in this study is proposed by aligning compatible sub-dimensions that reference analogous notions or concepts. This categorisation is grounded in the definitions of the sub-dimensions, the behavioural models they reference, and the contexts within which they are employed. However, this categorisation can suffer of limitations because of the interplay between dimensions and sub-dimensions. Indeed, for example, the concepts of extroversion and dominance are intimately linked to the valence and arousal, as encapsulated in the PAD model, suggesting that dominance could be considered as a component of extroversion. Similarly, trustworthiness is deeply intertwined with valence and arousal, and warmth significantly influences the competence (warmth and competence model).

In Sections 3.1 to 3.5, for each dimension separately, the influences of modalities and signals influence on these dimensions are

explained. Thus, it is now possible to highlight the main effect of a specific signal on a dimension. For instance, a smile, generally associated with positive valence [14], can lead to ambiguity if not clearly genuine [38], and its impact on friendliness can be reversed by contradictory gestures like authoritative arm movements [52]. Head movements such as nodding, typically signalling agreement [6, 9, 14, 18], can be perceived differently based on accompanying signals [14], highlighting the importance of signal congruence. Similarly, head tilts and orientation underscore the nuanced interpretation of emotional states, where a combination with other positive signals or specific contexts (like a forward-leaning posture) can significantly alter perceptions from negative to positive valence and arousal [6]. For verbal modalities, the linguistic choices made by IVAs, such as the use of inclusive pronouns and the level of formality, play pivotal roles in shaping perceived competence, friendliness and dominance [48]. Sentence length and content tone further influence perceptions of clarity, engagement, and emotions, affecting the user's interaction experience. Paralinguistic vocal immediacy such as pitch and speech rate stands out as a critical factor, with higher immediacy enhancing the IVA's anthropomorphism and friendliness [3]. Interruption handling times also markedly affect perceived friendliness and dominance, underscoring the delicate balance between responsiveness and assertiveness in verbal exchanges [17]. Overall, the synchronisation of verbal and non-verbal modalities, alongside the realism of vocal expressions, emerges as paramount in amplifying the IVA's perceived competence, trustworthiness, and naturalness in communication.

Throughout this paper, it is shown how the interplay and combination of various signals and modalities significantly influence the perception of IVAs, altering and amplifying user interpretations. For instance, a head shake stands out as a strongly negative cue [14], consistently associated with high arousal and negativity, irrespective of other modalities it accompanies. Contrarily, nodding typically signifies positive valence but can be perceived as neutral if conflicting signals accompany it [14]. Similarly, an IVA with a forward-leaning posture and head tilt can project positivity, a perception that might shift to neutrality or negativity when these signals are isolated. Interestingly, while a smile generally conveys friendliness, commanding gestures like finger-pointing can negate this effect [52]. The synchronisation of an IVA's gestures with its speech notably enhances perceived competence, highlighting the importance of congruence between verbal and non-verbal modalities. Moreover, the combination of voice and facial expressions, especially when aligned with body movements, serves as a powerful indicator of extroversion, underscoring the amplifying effect of multimodal communication. This synchronisation not only boosts the perception of competence [35] but also the persuasive power of the IVA, demonstrating the significant impact of integrated verbal and non-verbal modalities on user perceptions.

The interdependence between socio-affective perception and user characteristics should not be disregarded. Users' age has an influence on trust [23, 39] and on autonomy, persuasiveness and cooperativeness perception [43], with older users who are more likely to put their trust in the IVA. Divergent results are obtained in [47] regarding the impact of age on cooperativeness perception, since no correlation is established between these variables. With regard to competence [35], intelligence [43] and helpfulness [23],

no impact of age is identified. The gender of participants also plays a crucial role in the interaction dynamics with IVAs. Indeed, [7] shows that women tend to engage more closely and attentively with IVAs than men, who prefer a closer interpersonal distance particularly with female IVAs. Additionally, in [5], the authors show that the IVA, playing the role of a recruiter, is perceived significantly more believable by women than by men. Moreover, the gender of participants affects the perception of an IVA's friendliness, with female participants often perceiving IVAs as less friendly, regardless of the non-verbal behaviours displayed by the IVA [5, 7, 32, 47]. A last factor influencing the perception is the realism. Indeed, both animations quality [26, 29] and voice quality [16] crucially affect user perceptions of VA, with effective verbal and non-verbal behaviours enhancing believability.

In the discussion of the systematic review, it should be noted that only a small fraction of the studies specifically address immersive VR headsets. Out of the 32 papers analysed, merely five focus on immersive VR technologies. This observation underscores a significant gap in the literature, as immersive VR environments offer unique opportunities and challenges for the study of IVAs. These environments can potentially provide a more controlled and immersive context for examining the nuances of user interactions with IVAs, which might differ significantly from interactions in less immersive or monitor-based setups.

In addition, there is a significant body of research that is not included in this literature review. The strict adherence to the PRISMA method, while ensuring rigour, may have limited the scope of included studies and research groups considered and overlooked relevant research that falls outside the specified criteria. For example, the impact of cultural and demographic factors on socio-affective perceptions, though noted, warrants further exploration to understand how different cultural backgrounds influence user experiences with IVAs. It should be noted that several research works investigate these differences between cultures, comparing the perception of participants from Europe and Asia. Perception differs effectively when it comes to personality (i.e., friendliness) [22, 23, 32] and trustworthiness aspects [22]. The same statement can be made when it comes to social capabilities of IVAs. Indeed, users are more likely to take part in a conversation with agents when they belong to the same culture [31], highlighting a preference between users for signals attributed to their culture [30]. In [8], a comparison is established between Individualistic and Collectivist cultures. A significant impact of this cultural aspect is identified, influencing the perceived appropriateness of a discussion between virtual agents, either adopting a warmth and friendly behaviour, or a more aggressive and competitive conduct.

In conclusion, this systematic review sheds light on the intricate relationship between the multimodal behaviour of IVAs and the socio-affective perceptions of users, offering valuable insights and guidelines for the design of socio-affective IVAs. The findings underscore the need for a nuanced approach to IVA design that considers the full spectrum of non-verbal modalities and their interplay with verbal communication. As technology evolves and user expectations change, the field must continue to explore these dynamics, ensuring that IVAs remain effective, engaging, and capable of meeting the diverse needs of their users.

Table 1: Reasons and numbers of papers rejected during the PRISMA process.

Reason	Step 1	Step 2
Accessibility issue	0	4
Duplicate entry	1	1
Abstract or poster	2	0
Paper without DOI	11	18
No perception task	25	1
Perception of human agents	0	1
Perception of embodied avatars	9	2
Perception of robots	7	2
Faceless agents	0	5
Non-humanoid agents	2	2
No non-verbal or verbal behaviours	0	6
No monitor or VR	4	0
Participants with mean age under 18 years old	1	2
Non-Occidental participants	3	14
Pathological participants	8	0
Total	73	57

Table 2: Research articles considered in the systematic review

Reference	Dimensions	Sub-dimensions	Device	Non-verbal signals	Verbal signals	Nb of IVAs
Bee et al. [3]	Personality	Dominance, extroversion, friendliness	Monitor	Facial expressions, gaze, lip movements	Linguistic	1 (M)
Biancardi et al. [4]	Personality; Trustworthiness	Friendliness; Competence	Monitor	Arm movements, facial expressions	Linguistic	1 (F)
Callejas et al. [5]	Believability; Personality	Credibility; Dominance, friendliness	Monitor	Arm movements, facial expressions, gaze, head movements	Linguistic	1 (F)
Chollet and Scherer [6]	Emotions	Arousal, valence	Monitor	Facial expressions, gaze, gestures, head movements	/	1 (F/M) or 10 (5F, 5M)
Damian et al. [7]	Personality	Extroversion	Monitor	Gaze, gestures, group formation	/	2 (1F, 1M)
Demary et al. [9]	Emotions	Valence	Monitor	Facial expressions, gaze, body direction	/	1 (F/M)
Dermouche and Pelachaud [10]	Personality; Trustworthiness	Dominance, friendliness; Cooperativeness, helpfulness	Monitor	Arm movements, gaze, head movements, postures	/	1 (F)
Dermouche and Pelachaud [11]	Personality	Dominance, friendliness	Monitor	Arm movements, facial expressions, gaze, head movements, postures	/	3 (1F, 2M)
Etienne et al. [14]	Emotions	Arousal, valence	VR	Arm movements, head movements, gestures	/	1 (F/M)
Fourtoukidou et al. [16]	Believability; Emotions; Personality	Animacy, Anthropomorphism; Enthusiasm; Friendliness	Monitor	/	Paralinguistic	1 (M)
Gebhard et al. [17]	Personality	Dominance, friendliness, closeness	Monitor	/	Paralinguistic	1 (M)
Glemarec et al. [18]	Emotions	Arousal, valence	VR	Facial expressions, gaze, head movements, postures	/	10 (5F, 5M)
Hosseinpanah et al. [23]	Personality; Trustworthiness	Friendliness; Helpfulness, intelligence, trustworthiness	Monitor	Arm movements, facial expressions, head movements	/	1 (M)
Kang et al. [25]	Social capabilities	Intimacy	Monitor	Gaze, head movements	/	1 (M)
Kullmann et al. [26]	Believability	Naturalness, plausibility	VR	Gaze, facial expressions, postures	/	1 (M)
Lazzeri et al. [28]	Emotions	Arousal, valence	Monitor	Facial expressions	Linguistic, Paralinguistic	1 (F)
Lee et al. [29]	Believability	Believability, credibility	Monitor	Head movements	/	1 (M)
Marcoux et al. [32]	Personality	Friendliness	Monitor	Gaze, facial expressions, postures	/	1 (F/M)
Nguyen et al. [35]	Personality; Trustworthiness	Friendliness; Competence	Monitor	Gaze, gestures, facial expressions, postures	/	1 (F)
Nixon et al. [36]	Social capabilities	Social status	Monitor	Gaze, postures	/	1 (M)
Ochs et al. [38]	Emotions; Personality; Social capabilities; Trustworthiness	Valence; Friendliness; Mutual agreement, mutual understanding, stances; Epistemic stances	Monitor	Facial expressions	Linguistic	1 (F/M) or 2 (F)
Parmar et al. [39]	Believability; Trustworthiness	Animation quality; Persuasion, trustworthiness	Monitor	Animation fidelity, facial expressions, gestures	Linguistic, Paralinguistic	1 (F)
Podevin et al. [40]	Social capabilities	Intimacy, mutual comprehension	Monitor	Animacy, facial expressions, gaze, gestures, head movements, postures	Communication modality, linguistic, paralinguistic	2 (F)
Podevin et al. [41]	Social capabilities	Intimacy, mutual comprehension	Monitor	Facial expressions, gaze, gestures, head movements	Linguistic, paralinguistic	1 (F)
Prepin et al. [42]	Social capabilities	Mutual agreement, comprehension, understanding	Monitor	Facial expressions	/	2 (F)
Rosenthal-von der Pütten et al. [43]	Believability; Personality; Social capabilities; Trustworthiness	Believability, credibility; Dominance, friendliness; Mutual comprehension; Autonomy, competence, cooperativeness, intelligence, persuasion	Monitor	Gestures, gaze, sentence formulation	/	1 (M)
Sajjadi et al. [45]	Personality	Empathy, extroversion	VR	Facial expressions, gaze, gestures, postures	/	1 (F)
Straßmann et al. [47]	Personality; Trustworthiness	Dominance; Cooperativeness	Monitor	Arm movements, facial expressions, gaze, gestures, head movements, postures	/	1 (M)
von der Pütten et al. [48]	Social capabilities; Trustworthiness	Talkative, self-disclosure; Competence	Monitor	/	Linguistic, paralinguistic	1 (M)
Wang et al. [49]	Trustworthiness	Helpfulness, trustworthiness	Monitor	Facial expressions, gaze, head movements	/	1 (M)
Zojaji et al. [52]	Personality; Trustworthiness	Dominance, friendliness; Persuasion	Monitor	Arm movements, gaze	Linguistic, paralinguistic	8 (4M, 4F)
Zojaji et al. [53]	Personality; Trustworthiness	Dominance, friendliness; Persuasion	VR	Arm movements, gestures	Linguistic	8 (4M, 4F)

This table presents the selected papers after the second step of the PRISMA method (see Section 2).
For these papers, the dimensions and associated sub-dimensions presented in Section 3 are described.
The device, the verbal and non-verbal signals, the number of IVAs involved in these studies, and their genders are reported.
Note: M means Male and F means Female.

REFERENCES

[1] Elisabeth André, Martin Klesen, Patrick Gebhard, Steve Allen, and Thomas Rist. 1999. Integrating Models of Personality and Emotions into Lifelike Characters. In *Affective Interactions (IWAI 1999)*. Springer, 150–165. https://doi.org/10.1007/10720296_11

[2] Christoph Bartneck, Dana Kulić, Elizabeth Croft, and Susana Zoghbi. 2009. Measurement instruments for the anthropomorphism, animacy, likeability, perceived intelligence, and perceived safety of robots. *International Journal of Social Robotics* 1, 1 (2009), 71–81. https://doi.org/10.1007/s12369-008-0001-3

[3] Nikolaus Bee, Colin Pollock, Elisabeth André, and Marilyn Walker. 2010. Bossy or wimpy: Expressing social dominance by combining gaze and linguistic behaviors. In *10th International Conference on Intelligent Virtual Agents (IVA 2010)*. Springer, 265–271. https://doi.org/10.1007/978-3-642-15892-6_28

[4] Béatrice Biancardi, Chen Wang, Maurizio Mancini, Angelo Cafaro, Guillaume Chanel, and Catherine Pelachaud. 2019. A Computational Model for Managing Impressions of an Embodied Conversational Agent in Real-Time. In *8th International Conference on Affective Computing and Intelligent Interaction* (Cambridge, UK) *(ACII 2019)*. Institute of Electrical and Electronics Engineers Inc., 234 – 240. https://doi.org/10.1109/ACII.2019.8925495

[5] Zoraida Callejas, Brian Ravenet, Magalie Ochs, and Catherine Pelachaud. 2014. A Computational model of Social Attitudes for a Virtual Recruiter. In *13th International Conference on Autonomous Agents and Multiagent Systems* (Paris, France) *(AAMAS 2014)*. International Foundation for Autonomous Agents and Multiagent Systems, 93 – 100.

[6] Mathieu Chollet and Stefan Scherer. 2017. Perception of virtual audiences. *IEEE Computer Graphics and Applications* 37, 4 (2017), 50–59. https://doi.org/10.1109/MCG.2017.3271465

[7] Ionut Damian, Brigit Endrass, Peter Huber, Nikolaus Bee, and Elisabeth André. 2011. Individualized agent interactions. In *4th International Conference on Motion in Games (MIG 2011)*. 15–26. https://doi.org/10.1007/978-3-642-25090-3_2

[8] Nick Degens, Birgit Endrass, Gert Jan Hofstede, Adrie Beulens, and Elisabeth Andre. 2017. 'What I see is not what you get': why culture-specific behaviours for virtual characters should be user-tested across cultures. *AI and Society* 32, 1 (2017), 37–49. https://doi.org/10.1007/s00146-014-0567-2

[9] Guillaume Demary, Jean-Claude Martin, Stéphane Dubourdieu, Stéphane Travers, and Virginie Demulier. 2019. How do Leaders Perceive Stress and Followership from Nonverbal Behaviors Displayed by Virtual Followers?. In *19th ACM International Conference on Intelligent Virtual Agents* (Paris, France) *(IVA 2019)*. Association for Computing Machinery, 56–61.

[10] Soumia Dermouche and Catherine Pelachaud. 2018. Attitude modeling for virtual character based on temporal sequence mining: Extraction and evaluation. In *5th International Conference on Movement and Computing* (Genoa, Italy) *(MOCO 2018)*. Association for Computing Machinery. https://doi.org/10.1145/3212721.3212806

[11] Soumia Dermouche and Catherine Pelachaud. 2022. Leveraging the Dynamics of Non-Verbal Behaviors For Social Attitude Modeling. *IEEE Transactions on Affective Computing* 13, 2 (2022), 1072–1085. https://doi.org/10.1109/TAFFC.2020.2989262

[12] Tahirou Djara, Abdoul Matine Ousmane, and Antoine Vianou. 2019. Mood and personality influence on emotion. In *2nd International EAI Conference on Emerging Technologies for Developing Countries* (Cotonou, Benin) *(AFRICATEK 2018)*. Springer Verlag, 166–174. https://doi.org/10.1007/978-3-030-05198-3_15

[13] Paul Ekman and Wallace V Friesen. 1978. *Facial action coding system: Investigator's guide*. Consulting Psychologists Press.

[14] Elodie Etienne, Anne-Lise Leclercq, Angélique Remacle, Laurence Dessart, and Michaël Schyns. 2023. Perception of avatars nonverbal behaviors in virtual reality. *Psychology and Marketing* 40, 11 (2023), 2464–2481. https://doi.org/10.1002/mar.21871

[15] Susan T. Fiske, Amy J.C. Cuddy, and Peter Glick. 2007. Universal dimensions of social cognition: warmth and competence. *Trends in Cognitive Sciences* 11, 2 (2007), 77–83. https://doi.org/10.1016/j.tics.2006.11.005

[16] Sofia Fountoukidou, Uwe Matzat, Jaap Ham, and Cees Midden. 2019. Effects of a Virtual Model's Pitch and Speech Rate on Affective and Cognitive Learning. In *14th International Conference on Persuasive Technology* (Limassol, Cyprus) *(PERSUASIVE 2019)*. Springer Verlag, 16–27. https://doi.org/10.1007/978-3-030-17287-9_2

[17] Patrick Gebhard, Tanja Schneeberger, Gregor Mehlmann, Tobias Baur, and Elisabeth Andre. 2019. Designing the Impression of Social Agents' Real-time Interruption Handling. In *19th ACM International Conference on Intelligent Virtual Agents* (Paris, France) *(IVA 2019)*. Association for Computing Machinery, 19–21. https://doi.org/10.1145/3308532.3329435

[18] Yann Glemarec, Jean-Luc Lugrin, Anne-Gwenn Bosser, Aryana Collins Jackson, Cedric Buche, and Marc Erich Latoschik. 2021. Indifferent or Enthusiastic? Virtual Audiences Animation and Perception in Virtual Reality. *Frontiers in Virtual Reality* 2 (2021). https://doi.org/10.3389/frvir.2021.666232

[19] Jonathan Gratch, Anna Okhmatovskaia, Francois Lamothe, Stacy Marsella, Mathieu Morales, Rick J. Van Der Werf, and Louis-Philippe Morency. 2006. Virtual rapport. In *6th International Conference on Intelligent Virtual Agents (IVA 2006)*. Springer Verlag, 14–27. https://doi.org/10.1007/11821830_2

[20] Peter A. Hancock, Deborah R. Billings, Kristin E. Schaefer, Jessie Y. C. Chen, Ewart J. De Visser, and Raja Parasuraman. 2011. A Meta-Analysis of Factors Affecting Trust in Human-Robot Interaction. *Human Factors* 53, 5 (2011), 517–527. https://doi.org/10.1177/0018720811417254

[21] Willem K. B. Hofstee, Boele de Raad, and Lewis R. Goldberg. 1992. Integration of the Big Five and Circumplex Approaches to Trait Structure. *Journal of Personality and Social Psychology* 63, 1 (1992), 146–163. https://doi.org/10.1037/0022-3514.63.1.146

[22] Adineh Hosseinpanah and Nicole C. Krämer. 2021. Lost in Interpretation? The Role of Culture on Rating the Emotional Nonverbal Behaviors of a Virtual Agent. In *Cross-Cultural Design. Applications in Cultural Heritage, Tourism, Autonomous Vehicles, and Intelligent Agents*. Springer International Publishing, 350–368. https://doi.org/10.1007/978-3-030-77080-8_28

[23] Adineh Hosseinpanah, Nicole C. Krämer, and Carolin Straßmann. 2018. Empathy for Everyone? The Effect of Age When Evaluating a Virtual Agent. In *Proceedings of the 6th International Conference on Human-Agent Interaction* (Southampton, United Kingdom) *(HAI '18)*. Association for Computing Machinery, 184–190. https://doi.org/10.1145/3284432.3284442

[24] Ni Kang, Willem-Paul Brinkman, M Birna van Riemsdijk, and Mark A Neerincx. 2013. An expressive virtual audience with flexible behavioral styles. *IEEE Transactions on Affective Computing* 4, 4 (2013), 326–340. https://doi.org/10.1109/TAFFC.2013.2297104

[25] Sin-Hwa Kang, Jonathan Gratch, Candy Sidner, Ron Artstein, Lixing Huang, and Louis-Philippe Morency. 2012. Towards building a virtual counselor: Modeling nonverbal behavior during intimate self-disclosure. In *Proceedings of the 11th International Conference on Autonomous Agents and Multiagent Systems* (Valencia, Spain) *(AAMAS 2012, Vol. 1)*. International Foundation for Autonomous Agents and Multiagent Systems, 63–70.

[26] Peter Kullmann, Timo Menzel, Mario Botsch, and Marc Erich Latoschik. 2023. An Evaluation of Other-Avatar Facial Animation Methods for Social VR. In *Extended Abstracts of the 2023 CHI Conference on Human Factors in Computing Systems* (Hamburg, Germany) *(CHI EA '23)*. Association for Computing Machinery, Article 33, 7 pages. https://doi.org/10.1145/3544549.3585617

[27] Yunna Kwan, Sungwon Choi, Tae Rim Eom, and Tae Hui Kim. 2021. Development of a structured interview to explore interpersonal schema of older adults living alone based on autobiographical memory. *International Journal of Environmental Research and Public Health* 18, 5 (2021), 2316. https://doi.org/10.3390/ijerph18052316

[28] Nicole Lazzeri, Daniele Mazzei, Maher Ben Moussa, Nadia Magnenat-Thalmann, and Danilo De Rossi. 2018. The influence of dynamics and speech on understanding humanoid facial expressions. *International Journal of Advanced Robotic Systems* 15, 4 (2018). https://doi.org/10.1177/1729881418783158

[29] 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* (Toronto, Canada) *(AAMAS 2010, Vol. 1)*. International Foundation for Autonomous Agents and Multiagent Systems, 1257–1263.

[30] Birgit Lugrin, Elisabeth Andre, Matthias Rehm, Afia Lipi, and Yukiko Nakano. 2011. Culture-related differences in aspects of behavior for virtual characters across Germany and Japan. In *Proceedings of the 10th International Conference on Autonomous Agents and Multiagent Systems* (Taipei, Taiwan) *(AAMAS 2011, Vol. 1-3)*. Association for Computing Machinery, 441–448.

[31] Birgit Lugrin, Julian Frommel, and Elisabeth André. 2018. *Combining a Data-Driven and a Theory-Based Approach to Generate Culture-Dependent Behaviours for Virtual Characters*. Vol. 134. Springer International Publishing, 111–142. https://doi.org/10.1007/978-3-319-67024-9_6

[32] Audrey Marcoux, Marie-Hélène Tessier, and Philip L. Jackson. 2023. Nonverbal Markers of Empathy in Virtual Healthcare Professionals. In *Proceedings of the 23rd ACM International Conference on Intelligent Virtual Agents* (Würzburg, Germany) *(IVA '23)*. Association for Computing Machinery, Article 43, 4 pages. https://doi.org/10.1145/3570945.3607291

[33] Sébastien Mateo. 2020. A procedure for conduction of a successful literature review using the PRISMA method. *Kinesitherapie* 20, 226 (2020), 29–37. https://doi.org/10.1016/j.kine.2020.05.019

[34] Robert R. McCrae and Oliver P John. 1992. An introduction to the five-factor model and its applications. *Journal of Personality* 60, 2 (1992), 175–215. https://doi.org/10.1111/j.1467-6494.1992.tb00970.x

[35] 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 *Proceedings of the 15th International Conference on Intelligent Virtual Agents* (Delft, Netherlands) *(IVA '15)*. 167–180. https://doi.org/10.1007/978-3-319-21996-7_18

[36] Michael Nixon, Steve DiPaola, and Ulysses Bernardet. 2018. An Eye Gaze Model for Controlling the Display of Social Status in Believable Virtual Humans. In *IEEE Conference on Computational Intelligence and Games* (Maastricht, Netherlands) *(CIG 2018)*. IEEE, 8 pages. https://doi.org/10.1109/CIG.2018.8490373

[37] Magalie Ochs, Daniel Mestre, Grégoire De Montcheuil, Jean-Marie Pergandi, Jorane Saubesty, Evelyne Lombardo, Daniel Francon, and Philippe Blache. 2019.

Training doctors' social skills to break bad news: evaluation of the impact of virtual environment displays on the sense of presence. *Journal on Multimodal User Interfaces* 13 (2019), 41–51. https://doi.org/10.1007/s12193-018-0289-8

[38] Magalie Ochs, Catherine Pelachaud, and Ken Prepin. 2013. SOCIAL STANCES BY VIRTUAL SMILES. In *14th International Workshop on Image Analysis for Multimedia Interactive Services* (Paris, France) *(WIAMIS 2013)*. IEEE, 1 – 4. https://doi.org/10.1109/WIAMIS.2013.6616144

[39] Dhaval Parmar, Stefan Olafsson, Dina Utami, Prasanth Murali, and Timothy Bickmore. 2022. Designing empathic virtual agents: manipulating animation, voice, rendering, and empathy to create persuasive agents. *Autonomous Agents and Multi-Agent Systems* 36, Article 17 (2022). https://doi.org/10.1007/s10458-021-09539-1

[40] Delphine Potdevin, Céline Clavel, and Nicolas Sabouret. 2021. Virtual intimacy in human-embodied conversational agent interactions: the influence of multi-modality on its perception. *Journal on Multimodal User Interfaces* 15 (2021), 25–43. https://doi.org/10.1007/s12193-020-00337-9

[41] Delphine Potdevin, Celine Clavel, and Nicolas Sabouret. 2021. A virtual tourist counselor expressing intimacy behaviors: A new perspective to create emotion in visitors and offer them a better user experience? *International Journal of Human-Computer Studies* 150 (2021), 102612. https://doi.org/10.1016/j.ijhcs.2021.102612

[42] Ken Prepin, Magalie Ochs, and Catherine Pelachaud. 2013. Beyond backchannels: co-construction of dyadic stance by reciprocal reinforcement of smiles between virtual agents. In *Proceedings of the Annual Meeting of the Cognitive Science Society*. 1163–1168.

[43] Astrid M. Rosenthal-von der Pütten, Carolin Straßmann, Ramin Yaghoubzadeh, Stefan Kopp, and Nicole C. Krämer. 2019. Dominant and submissive nonverbal behavior of virtual agents and its effects on evaluation and negotiation outcome in different age groups. *Computers in Human Behavior* 90 (2019), 397–409. https://doi.org/10.1016/j.chb.2018.08.047

[44] James A. Russell and Albert Mehrabian. 1977. Evidence for a three-factor theory of emotions. *Journal of Research in Personality* 11, 3 (1977), 273–294. https://doi.org/10.1016/0092-6566(77)90037-X

[45] Pejman Sajjadi, Laura Hoffmann, Philipp Cimiano, and Stefan Kopp. 2018. On the Effect of a Personality-Driven ECA on Perceived Social Presence and Game Experience in VR. In *10th International Conference on Virtual Worlds and Games for Serious Applications (VS-Games)* (Würzburg, Germany). IEEE, Wurzburg, 1–8. https://doi.org/10.1109/VS-Games.2018.8493436

[46] Sinan Sonlu, Ugur Gudukbay, and Funda Durupinar. 2021. A Conversational Agent Framework with Multi-modal Personality Expression. *ACM Transactions on Graphics* 40, 1, Article 7 (2021), 16 pages. https://doi.org/10.1145/3439795

[47] Carolin Straßmann, Astrid M. Rosenthal-von der Pütten, Ramin Yaghoubzadeh, Raffael Kaminski, and Nicole C. Krämer. 2016. The effect of an intelligent virtual agent's nonverbal behavior with regard to dominance and cooperativity. In *Proceedings of the 16th International Conference on Intelligent Virtual Agents* (Los Angeles, USA) *(IVA '16)*. 15–28. https://doi.org/10.1007/978-3-319-47665-0_2

[48] Astrid M. von der Pütten, Laura Hoffmann, Jennifer Klatt, and Nicole C. Krämer. 2011. Quid pro quo? Reciprocal self-disclosure and communicative accomodation towards a virtual interviewer. In *Proceedings of the 10th International Conference on Intelligent Virtual Agents* (Reykjavik, Iceland) *(IVA '11)*. Springer-Verlag, 183–194. https://doi.org/10.1007/978-3-642-23974-8_20

[49] Isaac Wang, Rodrigo Calvo, Heting Wang, and Jaime Ruiz. 2023. Stop Copying Me: Evaluating nonverbal mimicry in embodied motivational agents. In *Proceedings of the 23rd ACM International Conference on Intelligent Virtual Agents* (Würzburg, Germany) *(IVA '23)*. Association for Computing Machinery, Article 49, 4 pages. https://doi.org/10.1145/3570945.3607322

[50] Jane Webster and Richard T. Watson. 2002. Analyzing the Past to Prepare for the Future: Writing a Literature Review. *MIS quarterly* 26, 2 (2002), xiii–xxiii. http://www.jstor.org/stable/4132319

[51] Zuhair Zafar, Ashita Ashok, and Karsten Berns. 2021. Personality Traits Assessment using P.A.D. Emotional Space in Human-robot Interaction. In *Proceedings of the 16th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications* (Vienna, Austria) *(VISIGRAPP 2021, Vol. 2)*. SciTePress, 111–118. https://doi.org/10.5220/0010161801110118

[52] Sahba Zojaji, Christopher Peters, and Catherine Pelachaud. 2020. Influence of virtual agent politeness behaviors on how users join small conversational groups. In *Proceedings of the 20th ACM International Conference on Intelligent Virtual Agents* (Virtual Event, Scotland, UK) *(IVA '20)*. Association for Computing Machinery, Article 59, 8 pages. https://doi.org/10.1145/3383652.3423917

[53] Sahba Zojaji, Adam Červeň, and Christopher Peters. 2023. Impact of Multimodal Communication on Persuasiveness and Perceived Politeness of Virtual Agents in Small Groups. In *Proceedings of the 23th ACM International Conference on Intelligent Virtual Agents* (Würzburg, Germany) *(IVA '23)*. Association for Computing Machinery, Article 18, 8 pages. https://doi.org/10.1145/3570945.3607356