Verbal Mimicry Predicts Social Distance and Social Attraction to an Outgroup Member in Virtual Reality


Document Version:
Peer reviewed version

Queen's University Belfast - Research Portal:
Link to publication record in Queen's University Belfast Research Portal

Publisher rights
2020 IEEE. This work is made available online in accordance with the publisher's policies. Please refer to any applicable terms of use of the publisher.

General rights
Copyright for the publications made accessible via the Queen's University Belfast Research Portal is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy
The Research Portal is Queen's institutional repository that provides access to Queen's research output. Every effort has been made to ensure that content in the Research Portal does not infringe any person's rights, or applicable UK laws. If you discover content in the Research Portal that you believe breaches copyright or violates any law, please contact openaccess@qub.ac.uk.
Verbal Mimicry Predicts Social Distance and Social Attraction to an Outgroup Member in Virtual Reality

Salvador Alvidrez  
School of Psychology  
Queen’s University Belfast  
Belfast, United Kingdom  
s.alvidrez@qub.ac.uk

Jorge Peña  
Department of Communication  
University of California at Davis  
Davis, United States  
jpena@ucdavis.edu

Abstract—The present study analyzes the extent to which verbal mimicry contributes to improving outgroup perceptions in virtual reality (VR) interactions. Particularly, this study examined the interplay between avatar customization, the salience of a common ingroup identity, and verbal mimicry in 54 VR dyads comprising users from different ethnic backgrounds. Participants were asked to customize their avatars to look either like themselves or someone completely different. Participants interacted wearing either similar avatar uniforms (salient common identity) or different clothes (non-salient identity). The linguistic style matching (LSM) algorithm was employed to calculate verbal mimicry in the communication exchanged during a joint task. The results suggested that verbal mimicry significantly predicted lesser social distance and greater social attraction towards the outgroup member. These results are discussed in terms of their contribution for potential intergroup models of avatar communication in immersive virtual environments (IVEs).

Keywords: mimicry; linguistic style matching; avatar customization; intergroup contact; virtual reality

I. INTRODUCTION

Mimicry, as peer-synchronized verbal behavior, has been related to positive social effects such as rapport, attraction, or group cohesiveness in both face-to-face (FtF) and text-based online interactions [1, 2]. People tend to converge in conversational styles or to mimic verbal and/or nonverbal features used by their counterparts for different social purposes, such as demonstrating empathy, agreeableness, or liking. Conversation partners may also choose to diverge if the one of them is disliked [3]. Sometimes these behavioral processes are displayed unconsciously in real life interactions, in which mimicry is almost undetectable for both speakers and external observers [4].

In online interactions, less attention has been paid to the verbal alignments emerging from avatar communication within IVEs, such as VR. In contrast to traditional communication venues, it has been suggested that embodying avatars in virtual reality influences the way in which users perceive themselves and others [5, 6]. This becomes even more relevant for interactions seeking to improve relations between members from different social groups (e.g., male/female, younger/older, White/Black, etc.), in which mimicry may play a key role type.

Technology-mediated contact between social groups has been increasingly portrayed as a feasible and effective alternative to FtF meetings, particularly for groups that differ in status, hold a large history of animosity, stigma, or that otherwise would never meet because of physical barriers or public restrictions [7, 8]. However, contact research in IVEs is relatively new and more scholarship is needed to predict optimal results in VR-based interactions, particularly for those interactions in which lasting identities such as race or ethnicity are ostensible.

This study originally tested a 2 (avatar customization: self vs. other) by 2 (common ingroup identity salience: salient vs. non-salient) between-groups experimental design on the dependent variable of social distance in a VR interaction [9]. The proposed model of indirect effects for this intervention included the mediating variables of engagement presence and VR group identification. Avatar customization has been linked to increases in presence (self, social) [10], and perceptions of “accountability” when taking part in a virtual group and the avatar mirrors the user’s physical traits [11]. The salience of common ingroup identities, on the other hand, has been proved to diminish intergroup bias as members focus more on the superordinate identity and less on their subgroup memberships [12]. Therefore, it was predicted that participants from different ethnic groups in the US (e.g., Asian, Caucasian, Black), interacting with participants from the Latino minority in VR, would feel less socially distant from this outgroup when a superordinate ingroup identity was salient (i.e., university campus), but only when their avatars were self-resembling. This effect would be mediated by the extent to which participants felt “present” in the conversation and identified with the VR group to which they were assigned. Contrary to expectations, findings suggested that those who customized their avatars in a way that mismatched their physical appearance felt more engaged into the conversation, and in turn, felt less socially distant from the target Latino outgroup. The salience of a common ingroup identity, at least in the original study, did not encourage participants to identify with their VR group.

Mimicry was not hypothesized in the original design, but it is introduced in this work as a post-hoc analysis with the purpose of outlining the potential role of avatar customization and common ingroup identity salience in producing mimicry, and the effect of mimicry on social
distance and social attraction. Self-categorization to a salient identity implies seeing the self and others as prototypical members of the category [13], leading to show ingroup favoritism and behave according to introjected norms like converging in a conversation with an ingroup member. Hence, we expect to see more mimicry when a common ingroup identity is made salient. Conversely, we expect to see more mimicry when avatars are customized to look different from the self, as users might feel less accountable (or more anonymous) if they say something inappropriate, which in turn may allow them to become more engaged into the conversation. Likewise, we expected to see less social distance and greater social attraction in dyads in which mimicry is high.

We used the LMS metric (14) to calculate verbal mimicry in interethnic dyads, which focuses on the use of function words such as personal pronouns, prepositions, or auxiliary verbs. Function words, compared to content words, are context-independent and have been proved to be useful in detecting associations between language and coordinated psychological states [15].

II. Method

A. Participants

A total of 135 students from undergraduate courses at a large public university in the United States participated in this study. They all received extra course credits in exchange for their participation in the study. Their ages ranged between 18 and 32 years; 87% of participants were female. 54.1% were Asian, 18.5% Hispanic, 18.5% Caucasian, and 9% were from “other” ethnic background. 80% of participants were in their senior year. The outgroup in this study is the Latino ethnic minority, therefore, the participation of Hispanic participants was dropped from the analysis in order to avoid ingroup bias. After filtering incomplete forms, low-quality recordings, and participants who failed to remember the ethnicity of their virtual partner, a final sample of 54 participations (dyads) was kept. As participants were randomly assigned to the experimental conditions, no significant differences were observed within groups in terms of gender ($\chi^2(1) = 0.01, p = 0.97$), ethnicity ($\chi^2(4) = 4.23, p = 0.37$) or number of participants in each condition prior to the experimental intervention ($F(1, 53) = 0.08, p = 0.77$).

B. Procedure

Participants registered for a lab session by filling out a pre-test questionnaire that included their demographic information. Upon arrival at the lab, an assistant greeted and instructed participants to use the VR equipment and meet another “student” from their same university in an immersive environment. The other student was actually a pre-trained research assistant or confederate. As dyads were gender-matched, confederate avatars were consistently portrayed as either “Carlos” or “María”; a 20-year old student in his/her senior year, born in Mexico and brought to the U.S. in his/her childhood. Their avatars included visual traits associated to the Latino communities: dark hair, skin, and eye color, as well as thick lips [16]. The appearance of the avatars used by the confederates can be seen in Fig. 1.

Participants were introduced to each other and brought to discuss what, in their opinion, are the most common problems faced by first-year university students in general, as well as to suggest solutions to tackle these problems. The participants completed a second questionnaire afterwards with several measures that included social distance ([17], $\alpha = 0.88$), social attraction ([18], $\alpha = 0.80$), engagement presence ([19], $\alpha = 0.80$), and VR group identification ([20], $\alpha = 0.80$). Then they were debriefed and thanked for their participation. Each session lasted 50 minutes in average and all sessions were digitally recorded.

C. Apparatus and Virtual Environment

An Oculus Rift DK2 head-mounted display (HDM) with integrated headphones and resolution of 2160x1200 pixels per eye was used. The HDM was connected to a desktop PC with an HDMI whereas a vertical sensor tracked the user’s head movements. A hand controller was also provided to participants.

We used vTime (www.vtime.net) as the VR platform for the immersive environment. vTime is a free social VR app that allows users to customize their avatars’ gender, skin color, age, shape, face, and outfit. It also offers pre-designed virtual scenarios and allows users to upload 360° photos in order to set background scenarios. vTime renders participants’ head movements onto their avatars in real-time. Users’ voices also trigger their avatars’ lip and hand movements and mannerisms in order to make conversations more natural. After entering the room, which allows the simultaneous interaction of a maximum of four users, avatars remain seated facing the other interaction partners.

D. Manipulations

For avatar customization, participants were instructed to customize the facial traits of their avatars (hair, eyes, skin, lips, nose, etc.), to make them look either like themselves or someone else. Having done this, the program asked the user to pose for a virtual camera that takes an instant picture and shows his/her customized avatar. In addition, in order to instill a sense of embodiment, a research assistant instructed participants to inspect the virtual environment and themselves, including looking down to their feet, where, instead of their real body, participants observed their avatar’s body from a first-person perspective.

![Fig. 1. Avatars used by the trained confederates. “Carlos” and “María”](image)
The salience of a common ingroup identity was manipulated by matching avatar outfits and verbally emphasizing belongingness to their university campus. The avatars used by the participant and the confederate wore clothes that mirrored their university’s institutional colors, and were told that they comprised a campus virtual group. In addition, other visual cues were included with the aim of increasing the salience of this institutional identity, such as portraying a 360° picture of the front entrance to the campus’ main library as a background. In contrast, avatars wore black and white colors when this identity was not salient, and a random country-side road was used as a background. Fig. 2 summarizes these experimental conditions.

E. Linguistic Analysis

In order to calculate LSM for all dyad interactions, we followed the protocol stipulated by Gonzales et al [14]. First, all verbal exchanges were transcribed and divided into two separate text files (participant/confederate). Then, both files were submitted to a Linguistic Inquiry and Word Count (LIWC) analysis, which calculates the percentage of total words in a text that can be assigned into nine function-word categories: personal pronouns (e.g., I, they, we), impersonal pronouns (e.g., it, that, anything), articles (e.g., an, the), conjunctions (e.g., and, but, because), prepositions (e.g., in, under, about), auxiliary verbs (shall, be, was), negations (no, not, never), common adverbs (e.g., hardly, often), and quantifiers (e.g., many, few). Function words are used as a measure of LSM because they are produced in a non-conscious manner and their interpretation is not linked to the context in which the conversation takes place. Function words create the scaffolding of the language, rather than conveying semantic meaning. Nouns or verbs, for instance, might be more ambiguous and difficult to interpret. Hence, function words can be measured across different domains of interaction (e.g., social, task) with relatively greater precision. The calculation of LSM scores for each category was performed using the following formula:

$$LSM_{conjs} = 1 - \frac{|conjs1 - conjs2|}{(conjs1 + conjs2)}$$

Accordingly, conjs1 is the percentage of conjunctions used by the confederate and conjs2 is the percentage used by the participant. The resulting score ranges between 0 and 1, with scores closer to 1 representing greater stylistic similarity (mimicry). The LSM scores for each category and a total score averaged from all nine categories are shown in Table 1.

F. Design

As mentioned before, we took the original 2 (avatar customization: self vs. other) by 2 (common ingroup identity: salient vs. non-salient) between-groups experimental design on the dependent variables of social distance and social attraction in a VR interaction. However, in order to understand if the experimental manipulations affect mimicry and how mimicry affects the dependent variables, first we analyze the influence of avatar customization and common ingroup identity salience on mimicry, and then the influence of mimicry on social distance and social attraction.

III. RESULTS

A. Previous Results

This study did not predict mimicry in its original design, as the focus was placed on the effects of avatar customization and salience of a common ingroup identity on social distance, and the mediating effects of engagement presence and VR group identification [see 9].

<table>
<thead>
<tr>
<th>Category</th>
<th>M</th>
<th>SD</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adverbs</td>
<td>0.92</td>
<td>0.05</td>
<td>0.76</td>
<td>0.99</td>
</tr>
<tr>
<td>Articles</td>
<td>0.88</td>
<td>0.08</td>
<td>0.67</td>
<td>0.98</td>
</tr>
<tr>
<td>Auxiliary verbs</td>
<td>0.92</td>
<td>0.07</td>
<td>0.66</td>
<td>0.99</td>
</tr>
<tr>
<td>Conjunctions</td>
<td>0.91</td>
<td>0.06</td>
<td>0.70</td>
<td>0.98</td>
</tr>
<tr>
<td>Impersonal pronouns</td>
<td>0.91</td>
<td>0.08</td>
<td>0.61</td>
<td>0.99</td>
</tr>
<tr>
<td>Negations</td>
<td>0.79</td>
<td>0.16</td>
<td>0.36</td>
<td>0.98</td>
</tr>
<tr>
<td>Personal pronouns</td>
<td>0.94</td>
<td>0.04</td>
<td>0.82</td>
<td>0.99</td>
</tr>
<tr>
<td>Prepositions</td>
<td>0.91</td>
<td>0.07</td>
<td>0.66</td>
<td>0.98</td>
</tr>
<tr>
<td>Quantifiers</td>
<td>0.90</td>
<td>0.08</td>
<td>0.65</td>
<td>0.99</td>
</tr>
<tr>
<td>Total LSM score</td>
<td>0.90</td>
<td>0.03</td>
<td>0.76</td>
<td>0.96</td>
</tr>
</tbody>
</table>

* Mean scores refer to the averaged dyad LSM score per each category. SD is the standard deviation of the mean LSM scores across all 54 dyads. Total LSM score is the average score obtained from aggregating all the function word categories. LSM = Linguistic style matching.
Manipulation checks showed that participants who embodied a self-resembling avatar felt significantly greater identification with their avatars ($M = 4.51$, $SD = 1.13$) than those who embodied an avatar resembling someone different ($M = 2.62$, $SD = 1.25$). Participants also identified more with their virtual group in those conditions in which a common ingroup identity was salient ($M = 6.07$, $SE = .84$), compared to when it was not ($M = 5.88$, $SE = .91$), although this difference was only marginal. The model of indirect effects showed a significant mediating effect of customizing self-resembling avatars on social distance by increasing engagement presence; this is, participants using self-resembling avatars perceived less social distance towards Latinos as long as they felt engaged with the conversation, but only when their uniforms did not match their partners’ and no mentions to the campus identity were made. In sum, the manipulation of a common ingroup identity (university campus) failed to make participants identify with the virtual group (i.e., campus VR group), and this failure was evident in subsequent analyses. As the predictions based on group membership failed to explain reductions in social distance, the source of the positive outcome was thought to have been produced during the interaction rather than as consequence of the experimental manipulations.

B. IVs Predicting Mimicry

For the current study we used the same dataset from our previous study, with the addition of the transcriptions made from the conversations in each dyad. Ordinary least squares (OLR) linear regressions were performed for all tests. For the first test, we regressed the LSM scores on avatar customization (self = 1, other = 0) and common ingroup identity (salient = 1, absent = 0). The participants’ sex and the amount of verbatim produced (word count) were also included in the regression in order to control their potential influence on the LSM scores. Contrary to expectations, neither a common ingroup identity ($b = -0.05$, $p = 0.57$) nor avatar customization ($b = 0.06$, $p = 0.50$) predicted the LSM scores. A second model including the interaction term between common ingroup identity and avatar customization was tested, but also failed to reach statistical significance. While sex did not predict LSM, word count significantly explained 24% of the variation in LSM ($R^2 = 0.24$, $p < 0.00$). In sum, the independent variables did not affect mimicry in the VR interaction. Table 2 shows a complete list of regression scores.

C. Mimicry Predicting DVs

The second test corresponds to the influence of mimicry on the dependent variables of social distance and social attraction. Social distance, which measures the extent to which respondents agree to items like “I accept having someone from the Latino minority as a member of my family”, was regressed on the LSM scores (higher scores = less distance). Engagement presence, VR group identification, sex, and word count were also added to the model to control their potential influence. Although the overall model did not reach significance, LSM alone significantly predicted social distance ($b = 11.85$, $p = 0.04$), whereas none of the other variables included in the model did so.

### TABLE II. IVs Predicting LSM

<table>
<thead>
<tr>
<th>Independent variablesa</th>
<th>Main effects model</th>
<th>Interaction model</th>
</tr>
</thead>
<tbody>
<tr>
<td>Common ingroup identity</td>
<td>$-0.00$</td>
<td>$-0.00$</td>
</tr>
<tr>
<td>Avatar customization</td>
<td>$0.00$</td>
<td>$0.00$</td>
</tr>
<tr>
<td>Sex</td>
<td>$-0.00$</td>
<td>$-0.00$</td>
</tr>
<tr>
<td>Word count</td>
<td>$2.18^{**}$</td>
<td>$2.16^{**}$</td>
</tr>
<tr>
<td>Interaction term common ingroup identity x avatar customization</td>
<td>$-0.00$</td>
<td>$0.01$</td>
</tr>
<tr>
<td>$R^2$</td>
<td>$0.24$</td>
<td>$0.24$</td>
</tr>
<tr>
<td>Adjusted $R^2$</td>
<td>$0.18$</td>
<td>$0.16$</td>
</tr>
</tbody>
</table>

a Common ingroup identity was coded as salient (1) or absent (0). Avatar customization was coded as either self-resembling (1) or other-resembling (0). Sex was coded as male (1) and female (2).

An additional regression was performed on social attraction, which unsurprisingly, showed that LSM positively predicts social attraction ($b = 12.41$, $p < 0.00$), after controlling for engagement presence ($b = 0.28$, $p < 0.00$), and VR group identification ($b = 0.16$, $p = 0.03$). Participants’ sex also predicted social attraction ($b = 0.68$, $p < 0.00$). The overall model explained 53% of the variation in attraction ($R^2 = 0.53$, $p < 0.00$). Table 2 shows a complete list of regression scores.

### IV. DISCUSSION

The primary goal of the present study was to test the role of verbal mimicry in outgroup perceptions in VR interactions. Verbal mimicry was calculated using the LSM algorithm, which analyzes verbal alignments in social interactions. Although the original 2x2 experimental design did not include mimicry, the IVs were tested first as potential predictors of LSM; however, mimicry was not affected by avatar customization and neither the salience of a common ingroup identity. As seen in our previous study [9], the salient institutional category was probably not relevant for participants in identifying with and influencing their verbal alignments.

### TABLE III. LMS Predicting DVs

<table>
<thead>
<tr>
<th>Independent variablesb</th>
<th>Social distance</th>
<th>Social attraction</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\beta$</td>
<td>$SE(\beta)$</td>
<td>$\beta$</td>
</tr>
<tr>
<td>LSM</td>
<td>$11.81^{*}$</td>
<td>$5.73$</td>
</tr>
<tr>
<td>Engagement presence</td>
<td>$0.06$</td>
<td>$0.16$</td>
</tr>
<tr>
<td>VR group identification</td>
<td>$0.04$</td>
<td>$0.17$</td>
</tr>
<tr>
<td>Sex</td>
<td>$-0.25$</td>
<td>$0.51$</td>
</tr>
<tr>
<td>Word count</td>
<td>$-4.37$</td>
<td>$0.00$</td>
</tr>
<tr>
<td>$R^2$</td>
<td>$0.10$</td>
<td>---</td>
</tr>
<tr>
<td>Adjusted $R^2$</td>
<td>$0.01$</td>
<td>---</td>
</tr>
</tbody>
</table>

b $p < 0.05$. **$p < 0.01$

---

a Social distance was measured on a 7-point scale. Social attraction refers to feeling that the confederate might be their friend. Engagement presence refers to feeling engaged with the virtual conversation. VR group identification refers to feeling identified with the VR campus dyad to which they were assigned.
Moreover, in contrast to previous interventions, matching avatar outfits (regardless of the institutional prime) did not trigger self-categorization or group prototypicality perceptions [21, 22]. Thus, no alignments in verbal exchanges based on group membership were observed. Likewise, the verbatim produced within each dyad, as the only significant predictor of mimicry, may have depended upon individual traits (e.g., extraversion) or the type of activity (e.g., social, task) rather than identifying with one’s own avatar or being accountable by self-resembling avatars. For example, Riordan, Markman, and Stewart [23] observed more textual alignment in online conversations when participants interacted in a social activity compared to a task activity.

The second round of tests examined the influence of mimicry on changes in outgroup perceptions. Results showed that mimicry alone significantly reduced social distance towards the Latino outgroup. In congruence with previous studies carried out in online settings, the groups in which verbal convergence was higher social distance was reported as being lowered. This also resonates with previous literature on intergroup relations: in multilingual societies, there is usually a veiled hierarchy in which the majority cultural group is placed at the top, while ethnic minorities are ranked closer or further away from this majority group on the basis of perceived cultural similarities [24]. In this case, the source of similarity came from conscious or unconscious verbal accommodation during a VR interaction, which also predicted social attraction significantly [15].

Nevertheless, it is intriguing that mimicry was the only significant predictor in the model of social distance. This means that, engagement presence and VR group identification, together with mimicry, may well predict social attraction but not social distance. In other words, being engaged in-or immersed into - the conversation with an outgroup member, as well as feeling part of the virtual group in which he or she is a member, is enough for the participant to feel socially attracted (e.g., friendly) to the confederate but not necessarily to feel socially closer to the whole outgroup (e.g., being part of the family). This resonates with earlier claims about the need for prolonged computer-mediated interactions in order to reach similar levels of intimacy, affinity, or trust as in F2F encounters [25]; particularly in intergroup settings [26]. However, it is yet to be seen if communication exchanges in 3D immersive environments such as VR require the same amount of time to achieve these outcomes compared to other traditional 2D virtual exchanges (e.g., Zoom). In any case, this study’s contributions are twofold: on one hand, it made clear that customizing avatars and interacting either in the presence or absence of a visually-induced group identity may not necessarily affect mimicry. In other words, personalizing avatars and prompting institutional affiliation do not automatically influence self-categorization to the salient group, and thus, to converge verbally or psychologically. On the other hand, according to the LSM analysis performed in this study, there is now empirical evidence that verbal mimicry actually emerges in avatar communication within an immersive VR setting. Nonetheless, the source of mimicry is less clear, at least in this study. A potential explanation might be that the topic discussed by dyads (problems faced by 1st year university students), probably triggered experiences familiar to participants (e.g., nostalgia, loneliness, anxiety), which in turn may have led them to converge unconsciously with the confederate’s speech. Future studies in this line may vary the topic of conversation and the nature of the interaction for corroborating these results.

There were different limitations to this study that deserve to be emphasized. First, the sample of participants was evidently eschewed towards female students, and although randomization of participants to conditions kept the ratio constant across groups, the participants’ sex played a relevant role in predicting social attraction, with higher scores among female participants. Future replies should gather a more balanced sample with regard to sex and ethnicity. In fact, future replies might want to examine the exclusive relationship between people from the cultural ethnic majority (i.e., White/European/Caucasian) and different ethnic minorities in order to verify whether social VR interactions work uniformly for all groups or not. Also, in light of new intriguing results derived from intergroup encounters in VR [27], it would be convenient to explore the extent to which embodying Latino avatars, for instance, would increase not only perspective-taking effects (i.e., feeling “in the shoes” of an outgroup member), but also processes of group influence led by visual cues to racial embodiment (e.g., reverse racial ingroup bias).

Another relevant limitation was related to the VR platform used in this study. We used a freely available system for social networking in VR (vTime), which allows users to hold immersive social interactions in real time. However, users are not in control of the avatar’s hand movements or its mobility, which may have decreased a sense of embodiment, agency, and telepresence as consequence. Because of that reason this study did not included an analysis of nonverbal mimicry, which seems mandatory for complementing our findings in verbal mimicry (see Hasler et al., [27]). Future research might incorporate full body haptic sensors together with eye-lip tracking technology in order to examine the influence of different degrees of embodiment and nonverbal behavior in intergroup encounters in VR.

ACKNOWLEDGMENT

The authors wish to thank Bianca Ocampo for her tireless support with the text analysis, as well as Kelly Torbet, Timmy Lui, Ian Johan-Gomez, Kimmiko James, Hannah Sarlatte, Abir Rahman, Emily Fong, Jiaying Wang, Michelle Tran, Cheng (Ethan) Ma, Robert Keene, Elizabeth Quezada, Daniel Gonzalez, Ivan Ornelas, Stephanie Pomales, Sara Geyer, Juliana Sierra Luna, Man Io Sou, and many other research assistants for their contribution to the original study.

REFERENCES


