How ECA vs Human Leaders Affect the Perception of Transactive Memory System (TMS) in a Team

Beatrice Biancardi*, Patrick O'Toole†, Ivan Giaccaglia*‡§,
Brian Ravenet§, Ian Pitt†, Maurizio Mancini¶ and Giovanna Varni*
*LTCI, Télécom Paris, Institut Polytechnique de Paris, France
†University College Cork, Ireland
‡University of Trento, Italy
§LISN-CNRS, University Paris-Saclay, France
§Sapienza University of Rome, Italy
Email: *name.surname@telecom-paris.fr
Email: †patrick.otoole@umail.ucc.ie, ianp@cs.ucc.ie
Email: ‡ivan.giaccaglia@gmail.com
Email: §brian.ravenet@universite-paris-saclay.fr
Email: ¶m.mancini@di.uniroma1.it

2021 9th International Conference on Affective Computing and Intelligent Interaction (ACII) | 978-1-6654-0019-0/21/$31.00 ©2021 IEEE | DOI: 10.1109/ACII52823.2021.9597454

***Abstract*—Transactive Memory System (TMS) is a mental representation of the distribution of knowledge between the members of a team. Can an Embodied Conversational Agent perform as well as a Human when intervening as a leader to support the development of the team's TMS? And, if yes, are there differences in the way the team perceives their respective interventions? In this paper, a perceptive online study is conducted on how Human leader interventions affect the perception of a team's TMS. The results are compared to the ones from a previous study evaluating an Embodied Conversational agent leader rather than a human one. Both the agent and the human adopt nonverbal behaviors characterizing 2 leadership styles: Transformational (TFL) and Transactional (TAL). TFL is expected to stimulate team members curiosity and creativity in problem-solving; instead, TAL emphasizes the role of the leader in supervising the team, providing it with feedback when needed. The results show that the intervention from both the agent and the human are perceived to potentially improve the perceived TMS of a team. Another interesting insight is that the TFL style works better when performed by the Human, where both the TAL and TFL style perform well when realized by the agent.**

***Index Terms*—Group Dynamics, Transactive Memory System, Embodied Conversational Agent, Team, Leadership**

I. Introduction

Human-Centered Computing (HCC) investigates how computational artifacts can support human activity, at the individual, team or society level [1], [2]. Latest progress in HCC and AI research is aimed at the development of intelligent computer systems able to improve and encourage the collaboration in teams, as reported in the "Research agenda on AI in team collaboration" [3]. The rationale behind this, is the shift in the role of the machine from being just a passive tool to becoming an effective teammate capable to engage human partners in a team collaboration [4].

The affective, behavioral and cognitive interactions between the members of a team, enable the emergence of team-specific social processes called "emergent states" (e.g., [5]). Among them, there is the Transactive Memory System (TMS), a meta-memory whose role is to store information about the specific knowledge owned by each member of a team.

As it emerges from several studies, the more machines, in the form of virtual agents or robots, exhibit a human-like embodiment and are capable to adhere to social norms, the more humans start to like and accept them as teammates [3], [6].

Embodied Conversational Agents (ECAs, that from now we call "agents") are computer interfaces displaying a graphical representation of a human-like virtual character [7]. Several studies investigated leadership, teams and agents, with the latter playing the role of followers, e.g., [8], [9]. Agents in the role of leaders were also studied, with the agent acting out different leadership styles in long distance interactions [10]. We are interested in investigating how an agent vs a Human teammate, playing the role of a leader with 2 different leadership styles, affects the perception of the TMS in a team of humans.

In this paper, we present an online[1] study, focused on an human leader, to inform the design of agent's interactions and understand the differences between agent and human interventions to create more realistic virtual teammates. Then, we compare our results with those in [11], which describes a similar study focused on an agent leader. In particular, we investigate how different leadership styles exhibited by an agent/human are perceived as facilitators of the TMS development in a working team.

II. Background

In this Section we introduce the psychological constructs we focus on in our study. A list of the abbreviations used in the paper with their definition can be found in Table I.

[1]The study was carried out online through videos and questionnaires, due to the the restrictions of the COVID-19 pandemics.

TABLE I
THE ABBREVIATIONS USED IN THIS PAPER.

Transactive Memory System	TMS	A mental representation of the distribution of knowledge between the members of a team.
Transformational Leadership	TFL	Leadership style stimulating followers' curiosity and creativity in problem solving. It tends to have a proactive way of influencing followers, using a supportive and collaborative approach.
Transactional Leadership	TAL	Leadership style emphasizing the role of the leader in supervising the team and providing it with feedback when needed. It tends to have a reactive way of influencing followers, using directive and dominant behaviors.

A. *Transactive Memory System*

Transactive Memory System (TMS) is defined as a team memory aimed at storing the knowledge owned by each member of a team. It emerges when team members use *"who knows what"* in the team to plan *"who will do what"* and results in: (i) a more efficient individual and collective performance [12]; (ii) a cooperative work division that allows a team to learn, remember and communicate the team's knowledge [13]; (iii) an improved knowledge transfer and retention of information [14]; an enhanced specialization of team members in different but compatible knowledge fields [15].

According to the literature, TMS is composed of 3 dimensions [16], [17]: *Specialization*, i.e., the differentiation of knowledge between the team members; *Credibility*, i.e., how much the members trust each other's knowledge; *Coordination*; i.e., the ability of the members to work together smoothly [18]. In this paper, aiming at improving team's performance and taking inspiration from previous research on the relationship between leadership and TMS [19], we focus on how different leadership styles displayed by an agent/human teammate affect the dimensions of TMS.

B. *Leadership Styles*

Several definitions of leadership exist, for example, according to Thompson: *"leadership is the ability to influence people to achieve the goals of a team"* [20]. Additionally, Forsyth identified some characteristics of this ability such as organizing, directing, coordinating, supporting and motivating followers [21]. Leaders, indeed, can coordinate people by giving clear cues for expertise and motivate the members of a team to recognize specific roles [22]. Specifically concerning TMS, previous work shows that leaders can exploit it to stimulate the emergence of interactions between members and helping them to recognize how to organize knowledge over different functions [23].

While different leadership styles are defined in the literature [24], the studies of Bachrach and colleagues [25] show that 2 types of leadership are positively associated with the development of TMS: Transformational Leadership (TFL) and Transactional Leadership (TAL).

TFL stimulates followers' curiosity and creativity in problem solving [26]. According to the literature, TFL is characterized by: articulating a vision (e.g., a common final goal), providing an appropriate model (e.g., leading by example), fostering the acceptance of team goals, having high performance expectations, providing individualized support, and promoting positive moods [27].

TAL instead, emphasizes the role of the leader in supervising the team and providing it with feedback when needed [28]. TAL leadership is particularly efficient to manage crises, by decreasing affective states such as panic, stress and anxiety in emergency scenarios [29].

According to Mackenzie et al., [30], TFL tends to have a more proactive way of influencing followers, whereas TAL tends to be more reactive. In general, TFL has a more supportive and cooperative approach with respect to TAL which is characterized by more directive and dominant behaviors. As Jung & Avolio state, TFL works deeply on the intrinsic values of the team, while TAL is more directed on intensifying the extrinsic values of the followers [31].

C. *Nonverbal communication in leadership*

Literature emphasizes the importance of nonverbal communication in the context of leadership [32]. In some particular cases, it can even be more important than verbal communication [33]. Previous work investigated what specific nonverbal behaviors are the most suitable to express TFL and TAL.

Awamleh & Gardner identify the following nonverbal behaviors related to TFL: maintaining eye contact, exhibiting vocal fluency (intended as a way of talking without exhibiting hesitations), using facial expressions (e.g., smiles), and engaging in dynamic hand and body gestures [34]. Furthermore, Schyns & Mohr report other nonverbal behaviors adopted by TFL leaders, such as: alternating between pacing and sitting on the edge of the desk, leaning towards the participant, and displaying rich facial expressions [35].

The literature on particular non-verbal behaviors characterizing TAL is not clearly reported. However, drawing on the idea of a direct and dominant leader, inspiration can be taken from works on dominant nonverbal behaviors [36]. For instance, dominant people generally adopt more downward palm gestures which are considered as a hint of something being wrong or a "power move" [37], [38] and they tend to display more closed and authoritarian postures (arms crossed) and neutral facial expressions.

D. *Embodied Conversational Agents and Leadership*

An Embodied Conversational Agent (ECA) is a Human-Machine interface designed to replicate human communicative abilities through the use of a virtual talking character [7]. In previous work investigating leadership, teams and agents, the latter mainly played the role of followers (e.g. [8], [9]). Jackson and colleagues investigated the role of an agent as a leader, acting different leadership styles, in very high distance interactions (space missions) focusing on the challenges coming from the latency of the communication [10]. Hayashi studied whether and how an agent can facilitate understanding and learning concepts during a collaborative peer-explanation activity [39]. This works shows that *"a conversational agent can facilitate a deeper understanding of concept when participants are attentive to its presence"*. To the best of our knowledge, however, no research was carried out to investigate whether an agent playing the role of a leader can impact the development of the TMS of a team.

III. Material & Method

To investigate how to design and implement teammate machines capable of developing and supporting TMS in teams, we needed to compare how people would react to the same intervention played by an agent vs a human teammate. Here we present a study focusing on a human teammate and we compare it with a previous one involving an agent teammate [11].

In both studies, we carried out an online perceptive study, where participants were asked to watch audio-visual stimuli and to provide their ratings on the TMS exhibited by a team before and after the intervention of a leader. The previous study in [11] followed the same methodology as the one presented in this paper (see Section III-A), the only difference being in the leader's embodiment (agent vs human).

A. Scenario & Stimuli

Each participant watched a sequence of stimuli composed as follows: an excerpt of the team acting in a scenario (Team Stimuli), followed by 2 excerpts of a leader displaying TFL or TAL leadership multimodal behaviors (Leader Stimuli).

The team acted in a design thinking scenario [40] in which 3 people, a psychologist, an environmental engineer and a designer, are involved in a design project to find solutions for improving green areas in urban environments, performing 3 activities. Previous work shows that design thinking is related to leadership styles and TMS: TFL and TAL are supportive to design thinking and useful to explain its benefits [41].

1) Team Stimuli

The 3 activities performed by the team were conceived to replicate steps that normally occur during design thinking and that also span the 3 dimensions of TMS. Moreover, the team members were instructed to make it clear through their behaviors that the team had an inability to leverage one of the 3 dimensions of TMS.

In the first activity, *"Division of Work"*, we emphasized the difficulty of the team to differentiate knowledge between its members (lack of Specialization). The members of the team discussed together and searched for topics related to green areas with the goal to acquire a broader knowledge on this issue.

In the second activity, *"Sharing Ideas"*, the team members did not show trust on each other's knowledge (lack of Credibility). Each member of the team presented their ideas about the project. Then, the psychologist and the engineer interact by themselves leaving the designer out, as he was having infeasible and useless ideas.

In the third activity, *"Choosing an Idea"*, the team members could not work together smoothly (lack of Coordination). Once they selected 3 ideas and discussed which one to implement, they started to change their mind to a different idea, resulting in a chaotic interaction. A frame from each activity is depicted in Figure 1.

The 3 Team Stimuli used in the presented study are the same ones exploited in [11]. They consist of three 1 minute long audio-video excerpts, each of which display the team members performing one of the above activities.

2) Leader Stimuli

We also created 6 Leader Stimuli for the leader (3 associated to TFL and the other 3 related to TAL nonverbal behaviors), designed to help the team to develop the dimension lacking in the corresponding activity excerpt.

In the previous study [11] the leader was played by an Embodied Conversational Agent, implemented through MARC, a publicly available agent platform [42]. In the present study, a professional actress could access the agent's performance exploited in [11], having the task of mimicking, as closely as possible, the expression of the complex socio-emotional behaviors of the agent through gestures, voice and facial expressions. The Leader Stimuli were recorded on a video-camera (30 fps, 1920x1080) and edited in Final Cut Pro, to clear up the sound and to optimize the lighting. The human actor's appearance was carefully made up to closely match the agent's one in [11] (see Table II). While the appearance of the actress could impact how she is perceived as a leader [43], it was not considered in the experimental design.

To design the facial expressions and the body movements of the agent/human actor, we relied on the literature mentioned in Section II-C. Despite the 2 leadership styles (TFL and TAL) sharing a common set of basic nonverbal behaviors (head nods and facial expressivity while talking), they were also characterized by some specific behaviors. The TFL-specific behaviors included: maintaining eye contact, smiling, dynamic gestures and open body postures. The TAL-specific behaviors included: authoritarian facial expressions, assertive palm downward gestures and closed body postures. Table II reports the agent and human actor's utterances used in each of the 3 activities for the 2 leadership styles (corresponding to the 6 Leader Stimulus).

B. Experimental Procedure

The present study[2] was carried out online through an *ad hoc* application developed on the `LimeSurvey` online survey tool[3]. Participants were mainly recruited via distribution on University mailing lists. Participation was voluntary and no information that would allow the identification or tracing of participants was collected (e.g. name, e-mail, IP address). Participants could withdraw from the study at any moment.

[2]The study was granted ethical approval by the Social Research Ethics Committee (SREC) of the University College Cork.

[3]https://www.limesurvey.org

(a) (b) (c)

Fig. 1. The 3 activities performed by the working team: (a) *"Division of Work"*, (b) *"Sharing Ideas"*, (c) *"Choosing an Idea"*.

TABLE II
THE AGENT AND HUMAN ACTOR'S UTTERANCES RELATED TO TFL AND TAL LEADERSHIP STYLES, PER TEAM ACTIVITY.

	Work Divison	**Sharing Ideas**	**Choosing an Idea**	**Agent**	**Human actor**
TFL	There are many ways this work could be divided. Could you please explain why you divided it this way?	I think you should all be inspired by the fact that you have different work approaches. Let's try to find some compromise together.	So, we need to find a way to reach a common agreement. Let's do a pros and cons table for each of these ideas, I'll help you to mediate the debate and make a decision.		
TAL	You must better know each other in the team. Tell us what could be your best contribution with this project considering your skills and interests.	There should be more trust between teammates. Look at these previous works I just sent you by email. It is possible to combine original design and well-established solutions.	You need to reach a common agreement. Write the pros and cons of each of the others' ideas. Then I want you to privately vote for an idea but you can't vote for yours. The idea that will get more votes wins.		

First, the participants were welcomed, given information on the study and how their data would be treated. If they gave their consent, participants were able to proceed. Next, the scenario and the characters acting in the Team Stimuli were presented. Then, they watched one of the Team Stimuli and filled out a questionnaire about its perceived TMS. After that, they watched 2 Leader Stimuli displaying TFL and TAL behaviors. After each of the Leader Stimuli, they filled out the same questionnaire again about the team TMS, which was pre-initialized with the ratings previously entered after the Team Stimulus. The order of the 3 team activities, as well as the Leader Stimuli were randomized between participants to avoid an affect of the ratings provided by participants. Finally, participants were thanked for taking part in the study and some demographic information (gender, age group and nationality) about them were collected.

Thirty-two people (17 males, 13 females, 2 N/A, mean age = 37y, SD=8y) participated in the study with the human actor, and twenty-height people (15 males, 13 females, mean age = 33y, SD=10y) in the study with the agent [11]. The two groups of participants were completely independent and were almost all European. We can thus consider that the two samples are almost matching in terms of age, sex and culture.

C. Independent and Dependent Variables

Since the study with the human actor followed the same methodology as the one presented in [11] (the only difference being in the leader's embodiment), this allows for comparing the results and merging the data in a mixed design. The independent variables manipulated are:

- *Activity*, understood as the 3 team activities, with levels: "Division of Work" (*Division*), "Sharing Ideas" (*Sharing*) and "Choosing an Idea" (*Choosing*);
- *Intervention*, understood as the type of intervention, if present, of the team leader, with levels: *Before*, *TFL*, *TAL*;

- *Leader Embodiment*, understood as the leadership type, with levels: *Agent*, *Human*.

The dependent variables are the 3 dimensions of TMS of the team as perceived by the participants: Specialization, Credibility, and Coordination. We measured them using items from a well-established questionnaire [17] already used in previous studies (e.g., [18], [44], [45]). We selected 3 items for each TMS dimension (for a total of 9 items), that participants were asked to rate on a 5-point Likert scale from *strongly disagree* to *strongly agree*. Table III reports the items we selected for each TMS dimension. These items were used to assess TMS after each activity as well as after each Leader Stimulus. Concerning the assessment of TMS after the Leader Stimulus, the questionnaire items were modified using future tense. This was because participants were asked to imagine how the behavior of the leader could affect the future behavior of the team.

D. Research Hypotheses

We address the following 4 research hypotheses (the first one focusing on the team, the other ones on the leader):
H1: in each activity, the team is perceived to have an inability to leverage the correlated TMS dimension (i.e., low scores of Specialization for the *Division* activity, of Credibility for the *Sharing* activity and of Coordination for the *Choosing* activity). This hypothesis aims to verify whether the stimuli we created were correctly perceived by the participants.
H2: the leader intervention positively affects the perceived TMS of the team (i.e., TMS scores for *TAL* and *TFL* interventions are higher in the *Before* level).
H3: *TAL* and *TFL* interventions have a different impact on the perceived TMS dimensions of the team.
H4: there is no effect of Leader Embodiment on the perception of team TMS (i.e., no effect of *Leader Embodiment* on TMS scores).

IV. Analysis and Results

Since Cronbach $\alpha \geq 0.73$ for each TMS dimension and each condition, we computed for each condition one score for Specialization, one for Credibility and one for Coordination by merging the 3 items of each dimension.

A. H1 - Activity

To test **H1**, we checked if each team activity was actually perceived with low scores for the correlated TMS dimension, compared to the other activities. In [11], **H1** was not validated for *Division*, partially validated for *Sharing* and completely validated for *Choosing*. In the following, we report the results of the same analysis conducted when using a *Human* as leader.

1) Specialization

We ran a one-way ANOVA with *Activity* as a within-subjects variable and Specialization as a dependent variable. The results show a main effect of *Activity*: $F(2, 62) = 9.8, p < 0.001$. Post-hoc pairwise comparisons with Bonferroni adjustment reveals that Specialization scores are significantly higher for *Division* ($M = 3.33, SD = 1.04$) compared to *Sharing* ($M = 2.62, SD = 1.18, t(32) = 3.72, p_{adj} < 0.001$) and *Choosing* ($M = 2.44, SD = 1.03, t(32) = 4.15, p_{adj} < 0.001$). No difference between *Sharing* and *Choosing* is found ($p_{adj} > 0.05$).

H1 is rejected for Specialization: the participants did not perceive lower Specialization in the team for the Division activity compared to the other team activities.

2) Credibility

Since the assumption of normality was not met for Credibility scores, we run a Friedman test with *Activity* as a within-subjects variable and Credibility as a dependent variable. The results show a main effect of *Activity*: $\chi^2(2) = 49, p < 0.001$, Kendall $W = 0.78$. Pairwise Wilcoxon signed-rank test reveals that Credibility scores for *Sharing* ($M = 2.16, SD = 1.09$) are significantly lower compared to *Division* ($M = 3.84, SD = 0.79, p_{adj} < 0.001, r = 0.81$), and significantly higher compared to *Choosing* ($M = 1.32, SD = 0.46, p < 0.01, r = 0.87$).

H1 is partially accepted for Credibility: the participants perceived lower Credibility in the team for the Sharing activity compared to the Division activity, but higher compared to the Choosing activity.

3) Coordination

Since the assumption of normality was not met for Coordination scores, we run a Friedman test, with *Activity* as a within-subjects variable and Coordination as a dependent variable. The results show that Coordination scores are significantly different across each *Activity*: $\chi^2(2) = 56.1$, Kendall's $W = 0.88, p < 0.001$. Pairwise Wilcoxon signed-rank test reveals that Coordination scores are significantly lower for *Choosing* ($M = 1.31, SD = 0.52$) compared to *Divison* ($M = 4.28, SD = 0.94, p_{adj} < 0.001, r = 0.87$), and *Sharing* ($M = 2.73, SD = 1.12, p_{adj} < 0.001, r = 0.85$).

H1 is fully accepted for Coordination: the participants perceived lower Coordination in the team for the Choosing activity, compared to the other activities.

B. H2, H3, H4 - Intervention and Leader Embodiment

1) Specialization

We run a 2x3x3 mixed ANOVA, with *Leader Embodiment* as a between-subjects variable, *Activity* and *Intervention* as within-subjects variables and Specialization as a dependent variable. Results show a main effect of *Activity* ($F(2, 116) = 27.16, p < 0.001$) and a main effect of *Intervention* ($F(1.76, 102.01) = 46.86, p < 0.001$). In addition, a two-way interaction is found between *Leader Embodiment* and *Intervention* ($F(1.76, 102.01 = 3.45, p < 0.05$) and between *Activity* and *Leader Embodiment* ($F(4, 232) = 7.83, p < 0.001$). No main effect of *Leader Embodiment* is found, supporting **H4**.

A one-way ANOVA reveals that the main effect of *Activity* is significant for every level of the other variables (all $p < 0.05$), except for the *TFL* intervention for the level *Human* ($p = 0.2$). In general, Specialization scores are higher for *Division* ($M = 3.82, SD = 0.99$) compared to the other activities

TABLE III
THE ITEMS OF THE QUESTIONNAIRE [17] USED TO ASSESS TMS.

TMS dimension	Selected Items
Specialization	Different team members were responsible for expertise in different areas. The specialized knowledge of several team members was needed to complete the project deliverables. Each team member knew which team member had expertise in specific areas.
Credibility	Team members were comfortable accepting procedural suggestions from the other team members. They were confident relying on the information that other team members brought to the discussion. They did not have much faith in other members' expertise.
Coordination	The team worked together in a well-coordinated fashion. The team needed to backtrack and start over a lot. There was much confusion about how they would accomplish a task.

($M = 3.19, SD = 1.13$ for *Sharing*; $M = 3.07, SD = 1.17$ for *Choosing*), which is in line with what we found while testing **H1**. Pairwise comparisons with Bonferroni correction show that this difference is significant for every condition (except for the one mentioned before, all $p_{adj} < 0.05$), while there is no difference between *Sharing* and *Choosing* (all $p_{adj} > 0.4$).

A one-way ANOVA reveals that the main effect of *Intervention* is significant for every level of the other variables (all $p < 0.001$). In general, Specialization scores increase after the intervention, so the scores given for *Before* ($M = 2.87, SD = 1.17$) are lower than those given after *TAL* ($M = 3.55, SD = 1.12$) or *TFL* ($M = 3.66, SD = 0.98$). **H2** is thus validated for Specialization.

Regarding the interaction effect between *Leader Embodiment* and *Intervention*, a two-way ANOVA reveals that this interaction is significant only for the *Sharing* activity ($F(2, 116) = 3.94, p < 0.05$).

Regarding the interaction effect between *Activity* and *Intervention*, a two-way ANOVA reveals that this interaction is significant only for the level *Human* ($F(4, 124) = 6.52, p < 0.001$).

Pairwise comparisons with Bonferroni correction show that for the level *Agent* of variable *Leader Embodiment*, the *TFL* intervention do not improve Specialization in the *Division* activity ($p_{adj} > 0.05$), while, for the level *Human*, *TAL* intervention do not improve Specialization in the *Sharing* activity ($p_{adj} > 0.05$). This seems to indicate that the TFL strategy works better when used by a human than an agent, and the TAL strategy better works when used by an agent than a human. TFL also perform better than TAL in the *Choosing* activity for the level *Human* ($t(32) = 2.68, p_{adj} < 0.05$). **H3** is thus partially validated for Specialization.

To summarize: *H2 and H4 can be accepted, whereas H3 can be only partially accepted. In particular, results show that, to improve a team's Specialization, TFL strategy seems to work better when performed by a human and TAL strategy seems to work better when performed by an agent, but in general the intervention of an agent has a similar effect on Specialization as seen with the intervention of a human*.

2) Credibility

We run a 2x3x3 mixed ANOVA, with *Leader style* as a between-subjects variable, *Activity* and *Intervention* as within-subjects variables and Credibility as a dependent variable. The results show a main effect of *Activity* ($F(2, 116) = 124.5, p < 0.001$) and a main effect of *Intervention* ($F(1.7, 98.35) = 54.28, p < 0.001$). In addition, a two-way interaction is found between *Activity* and *Intervention* ($F(4, 232) = 21.74, p < 0.001$). No main effect of *Leader Embodiment* is found, supporting **H4**.

A one-way ANOVA reveals that the main effect of *Activity* is significant for every level of the other variables (all $p < 0.001$). In general, Credibility scores for *Sharing* ($M = 2.75, SD = 1.15$) are lower compared to *Division* ($M = 3.98, SD = 0.88$) and higher compared to *Choosing* ($M = 2.27, SD = 1.11$), which is in line with what we found while testing **H1**. Pairwise comparisons with Bonferroni correction show that these differences are significant for every condition, except for 2 cases where Credibility scores are not significantly different between *Sharing* and *Choosing* (for *TAL* and *TFL* intervention for the level *Agent*, $p_{adj} > 0.08$).

A one-way ANOVA reveals that the main effect of *Intervention* is significant for every level of the other variables (all $p < 0.001$), except for *Division* activity ($p > 0.3$). In the other cases, Credibility scores increase after the intervention, so the scores given for level *Before* of variable *Intervention* ($M = 2.48, SD = 1.38$) are lower than those given after *TAL* ($M = 3.16, SD = 1.18$) or *TFL* ($M = 3.36, SD = 1.04$). **H2** is thus partially validated for Credibility.

Regarding the interaction effect between *Activity* and *Intervention*, a two-way ANOVA reveals that this interaction is significant for both types of leader embodiment (all $p < 0.001$).

Pairwise comparisons with Bonferroni correction show that (except for *Division* activity, where we already said that Credibility scores do not differ across *Intervention*) *TFL* gets higher scores than *TAL* for the level *Human* ($p_{adj} < 0.05$). Conversely, for the level *Agent* the two leadership strategies both increase Credibility scores compared to level *Before*, but do not significantly differ from each other ($p_{adj} > 0.5$). This is in line with the higher performance of TFL when performed by a human that we found for the Specialization scores. **H3** is thus partially validated for Credibility.

To summarize: *H4 can be accepted, whereas H2 and H3 can be only partially accepted. In particular, results show that, to improve team's Credibility, the TFL strategy seems to work better when performed by a human, while there is no difference between TFL and TAL in the case of an agent*.

3) Coordination

We run a 2x3x3 mixed ANOVA, with *Leader Embodiment* as a between-subjects variable, *Activity* and *Intervention* as within-subjects variables and Coordination as a dependent variable. The results show a main effect of *Activity* ($F(2, 116) = 110.32, p < 0.001$) and a main effect of *Intervention* ($F(2, 116) = 54.85, p < 0.001$). In addition, a two-way interaction is found between *Leader Embodiment* and *Intervention* ($F(2, 116) = 3.42, p < 0.05$) and between *Activity* and *Intervention* ($F(4, 232) = 58.78, p < 0.001$). No main effect of *Leader Embodiment* is found, supporting **H4**.

A one-way ANOVA reveals that the main effect of *Activity* is significant for every level of the other variables (all $p < 0.01$). In general, Coordination scores are lowest for *Choosing* ($M = 2.55, SD = 1.26$) compared to the other activities ($M = 3, SD = 1.09$ for *Sharing*; $M = 4.12, SD = 1.06$ for *Division*), which is in line with what we found while testing **H1**. Pairwise comparisons with Bonferroni correction show that this difference is significant for every condition, except for *TAL* and *TFL* interventions for both groups, where the Coordination scores are equally lower for *Sharing* and *Choosing* (all $p_{adj} > 0.3$).

A one-way ANOVA reveals that the main effect of *Intervention* is significant for every level of the other variables (all $p < 0.001$), except for *Division* activity ($p > 0.9$). In the other cases, Coordination scores increase after the intervention, so the scores given for *Before* ($M = 2.69, SD = 1.51$) are lower than those given after *TAL* ($M = 3.41, SD = 1.13$) or *TFL* ($M = 3.57, SD = 1.11$). **H2** is thus partially validated for Coordination.

Regarding the interaction effect between *Leader Embodiment* and *Intervention*, a two-way ANOVA reveals that this interaction is significant only for the *Choosing* activity ($F(2, 116) = 3.34, p < 0.05$).

Regarding the interaction effect between *Activity* and *Intervention*, a two-way ANOVA reveals that this interaction is significant only for both groups (all $p < 0.001$).

Pairwise comparisons with Bonferroni correction show that (except for *Division* activity, where we already said that Coordination scores do not differ across *Intervention*) *TFL* strategy gets higher scores than *TAL* for the level *Human* (all $p_{adj} < 0.05$). Conversely, for the level *Agent* the two leadership strategies both increase Coordination scores compared to level *Before*, but do not significantly differ to each other ($p_{adj} > 0.4$). **H3** is thus partially validated for Coordination.

To summarize: *H4 can be accepted, whereas H2 and H3 can be only partially accepted. In particular, results show that, to improve team's Coordination, TFL strategy seems to work better when performed by a human, while there is no difference between TFL and TAL in the case of an agent.*

V. Conclusion and Future Work

In this paper, we presented a perceptive online study about the impact of different leadership strategies on a team's development of Transactive Memory System (TMS). We compared the effect of such strategies when performed by an Embodied Conversational Agent and a human. We designed the behaviors of an agent and a human actor that characterize 2 leadership styles: Transformational (TFL) vs Transactional (TAL).

Results show that the intervention of both the agent and the human leader could improve the perceived TMS of the team. This could be explained by the fact that both the leaders are anthropomorphic, which is in line with [6]. Further studies are needed to explore different types of embodiment (e.g., a robot) and lower levels of anthropomorphism (e.g., a vocal assistant). Also, it was demonstrated that a female agent is perceived as more trustworthy than a male agent [46]: again, further investigation is needed in this direction. When looking in detail at our results, we observe that the TFL leadership strategy works better when performed by a human, while both TFL and TAL strategies generally impact the perception of TMS when they are performed by an agent. This is interesting since it would mean that when using a virtual leader, one could use different leadership styles without reducing the positive impact it could have on the team's TMS.

The team stimuli were conceived to exhibit a lack in one of the 3 TMS dimensions. The perception of these stimuli differed according to the TMS dimension: the activity that was best recognized was *Choosing an Idea* (lack of Coordination), *Sharing Ideas* (lack of Credibility) was partially recognized, while *Division of Work* (lack in Specialization) was not perceived as we expected. This limits the scope of interpretation we can do regarding this particular dimension and how it was perceived after each leaders' intervention. We are considering carrying out other experiments in a physical, instead of remote, environment, to investigate if the perception of the team TMS improves. This should also offer an environment in which we can more easily design an interaction scenario where the leader (an agent or a human) would be part of the team and not only an external observer.

In summary, this paper shows promising results supporting the use of Embodied Conversational Agents as effective team leaders, capable of engaging groups of users and supporting them in the development of complex affective and cognitive phenomena, such as TMS. After assessing the potential of Embodied Conversational Agents as leaders in a first online study, and comparing it with human leaders in the current one which replicated the design of the first one, our next step will be to leverage from these results to study and investigate their impact on teams' TMS in real-time interactive scenarios.

Acknowledgements

We would like to thank Sarah O'Toole for performing the human actor in our study. The work of I. Giaccaglia was funded by the DATAIA convergence institute "Programme d'Investissement d'Avenir" (ANR-17-CONV-0003). The work of P. O'Toole was supported by a Grant from Science Foundation Ireland under Grant number 18/CRT/6222.

REFERENCES

[1] E. Clarkson, J. A. Day, and J. D. Foley, "An educational digital library for human-centered computing," in *CHI'06 Extended Abstracts on Human Factors in Computing Systems*, 2006, pp. 646–651.

[2] J. Canny, "Human-centered computing (technical report). berkeley," *CA: University of California, Berkeley*, 2001.

[3] I. Seeber, E. Bittner, R. O. Briggs, T. De Vreede, G.-J. De Vreede, A. Elkins, R. Maier, A. B. Merz, S. Oeste-Reiß, N. Randrup *et al.*, "Machines as teammates: A research agenda on ai in team collaboration," *Information & management*, vol. 57, no. 2, p. 103174, 2020.

[4] T. W. Malone, "How human-computer'superminds' are redefining the future of work," *MIT Sloan Management Review*, vol. 59, no. 4, pp. 34–41, 2018.

[5] S. W. Kozlowski, "Advancing research on team process dynamics: Theoretical, methodological, and measurement considerations," *Organizational Psychology Review*, vol. 5, no. 4, pp. 270–299, 2015.

[6] J. Złotowski, D. Proudfoot, K. Yogeeswaran, and C. Bartneck, "Anthropomorphism: opportunities and challenges in human–robot interaction," *International journal of social robotics*, vol. 7, no. 3, pp. 347–360, 2015.

[7] J. Cassell, J. Sullivan, E. Churchill, and S. Prevost, *Embodied conversational agents*. MIT press, 2000.

[8] G. Demary, J.-C. Martin, S. Dubourdieu, S. Travers, and V. Demulier, "How do leaders perceive stress and followership from nonverbal behaviors displayed by virtual followers?" in *Proceedings of the 19th ACM International Conference on Intelligent Virtual Agents*, 2019, pp. 56–61.

[9] G. Demary, "Évaluation cognitive du leader dans une dyade hiérarchique: des comportements non verbaux du suiveur aux comportements de leadership," Ph.D. dissertation, 2018.

[10] A. C. Jackson, E. Bevacqua, P. D. Loor, and R. Querrec, "Modelling an embodied conversational agent for remote and isolated caregivers on leadership styles," *IVA 2019 - Proceedings of the 19th ACM International Conference on Intelligent Virtual Agents*, pp. 256–259, 2019.

[11] B. Biancardi, I. Giaccaglia, B. Ravenet, and G. Varni, "Virtual leaders supporting the development of transactive memory systems," in *Proceedings of the "32e Conférence Francophone sur l'Interaction Homme-Machine"*, 2021.

[12] N. Gupta and A. B. Hollingshead, "Differentiated versus integrated transactive memory effectiveness: It depends on the task." *Group Dynamics: Theory, Research, and Practice*, vol. 14, no. 4, p. 384, 2010.

[13] D. M. Wegner, "Transactive Memory: A Contemporary Analysis of the Group Mind," *Theories of Group Behavior*, pp. 185–208, 1987.

[14] L. Argote and P. Ingram, "Knowledge transfer: A basis for competitive advantage in firms," *Organizational behavior and human decision processes*, vol. 82, no. 1, pp. 150–169, 2000.

[15] S. W. Kozlowski and D. R. Ilgen, "Enhancing the effectiveness of work groups and teams," *Psychological science in the public interest*, vol. 7, no. 3, pp. 77–124, 2006.

[16] R. L. Moreland and L. Myaskovsky, "Exploring the performance benefits of group training: Transactive memory or improved communication?" *Organizational behavior and human decision processes*, vol. 82, no. 1, pp. 117–133, 2000.

[17] K. Lewis, "Measuring transactive memory systems in the field: Scale development and validation," *Journal of Applied Psychology*, vol. 88, no. 4, pp. 587–604, 2003.

[18] Y. Ren and L. Argote, "Transactive memory systems 1985-2010: An integrative framework of key dimensions, antecedents, and consequences," *Academy of Management Annals*, vol. 5, no. 1, pp. 189–229, 2011.

[19] V. Peltokorpi, "Transactive Memory Systems," *Review of General Psychology*, vol. 12, no. 4, pp. 378–394, 2008.

[20] L. L. Thompson and M. Thompson, "Making the team: A guide for managers," 2008.

[21] D. R. Forsyth, *Group dynamics*. Cengage Learning, 2018.

[22] J. R. Larson Jr, P. G. Foster-Fishman, and T. M. Franz, "Leadership style and the discussion of shared and unshared information in decision-making groups," *Personality and Social Psychology Bulletin*, vol. 24, no. 5, pp. 482–495, 1998.

[23] M. Soekijad, B. van den Hooff, M. Agterberg, and M. Huysman, "Leading to learn in networks of practice: Two leadership strategies," *Organization Studies*, vol. 32, no. 8, pp. 1005–1027, 2011.

[24] T. Nanjundeswaraswamy and D. Swamy, "Leadership styles," *Advances in management*, vol. 7, no. 2, p. 57, 2014.

[25] D. G. Bachrach and R. Mullins, "A dual-process contingency model of leadership, transactive memory systems and team performance," *Journal of Business Research*, vol. 96, no. November 2018, pp. 297–308, 2019. [Online]. Available: https://doi.org/10.1016/j.jbusres.2018.11.029

[26] E. Brymer and T. Gray, "Effective leadership: Transformational or transactional?" *Journal of Outdoor and Environmental Education*, vol. 10, no. 2, pp. 13–19, 2006.

[27] P. M. Podsakoff, S. B. MacKenzie, R. H. Moorman, and R. Fetter, "Transformational leader behaviors and their effects on followers' trust in leader, satisfaction, and organizational citizenship behaviors," *The leadership quarterly*, vol. 1, no. 2, pp. 107–142, 1990.

[28] M. Ljungblom, "A Comparative Study Between Developmental Leadership and Lean Leadership – Similarities and Differencies," *Management and Production Engineering Review*, vol. 3, no. 4, pp. 54–68, 2012.

[29] B. M. Bass, "Transformational leadership: industrial," *Military, and Educational Impact, Mahwah*, 1998.

[30] S. B. Mackenzie, P. M. Podsakoff, and G. A. Rich, "Transformational and Transactional Leadership and Salesperson Performance The Dangers of Poor Construct Conceptualization View project," *Journal of the Academy of Marketing Science*, vol. 29, no. 2, pp. 115–134, 2001.

[31] D. I. Jung and B. J. Avolio, "Opening the black box: An experimental investigation of the mediating effects of trust and value congruence on transformational and transactional leadership," *Journal of organizational Behavior*, vol. 21, no. 8, pp. 949–964, 2000.

[32] A. Darioly and M. S. Mast, "The role of nonverbal behavior in leadership: An integrative review." 2014.

[33] M. Remland, "Developing leadership skills in nonverbal communication: A situational perspective," *The Journal of Business Communication (1973)*, vol. 18, no. 3, pp. 17–29, 1981.

[34] R. Awamleh and W. L. Gardner, "Perceptions of leader charisma and effectiveness: The effects of vision content, vision delivery, and organizational performance," in *Proceedings of the Southern Management Association Meetings*, 1997, pp. 76–78.

[35] B. Schyns and G. Mohr, "Nonverbal elements of leadership behaviour," *German Journal of Human Resource Management*, vol. 18, no. 3, pp. 289–305, 2004.

[36] J. K. Burgoon and N. E. Dunbar, "Nonverbal expressions of dominance and power in human relationships," *The Sage handbook of nonverbal communication*, vol. 2, pp. 279–297, 2006.

[37] A. Kendon, "Language and gesture: Unity or duality," *Language and gesture*, vol. 2, 2000.

[38] G. Imai, "Gestures: Body language and nonverbal communication.[web document]," 1996.

[39] Y. Hayashi, "On pedagogical effects of learner-support agents in collaborative interaction," *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 7315 LNCS, pp. 22–32, 2012.

[40] D. Pusca and D. O. Northwood, "Design thinking and its application to problem solving," *Global Journal of Engineering Education*, vol. 20, no. 1, p. 3, 2018.

[41] J. Schweitzer, "Transformational Leadership , Design Thinking and the Innovative Firm," 2014.

[42] M. Courgeon and C. Clavel, "MARC: A framework that features emotion models for facial animation during human-computer interaction," *Journal on Multimodal User Interfaces*, vol. 7, no. 4, pp. 311–319, 2013.

[43] S. Sczesny and U. Kühnen, "Meta-cognition about biological sex and gender-stereotypic physical appearance: Consequences for the assessment of leadership competence," *Personality and Social Psychology Bulletin*, vol. 30, no. 1, pp. 13–21, 2004.

[44] K. Kwon and D. Cho, "How transactive memory systems relate to organizational innovation: the mediating role of developmental leadership," *Journal of Knowledge Management*, vol. 20, no. 5, pp. 1025–1044, 2016.

[45] P. Marques-Quinteiro, L. Curral, A. M. Passos, and K. Lewis, "And now what do we do? The role of transactive memory systems and task coordination in action teams," *Group Dynamics*, vol. 17, no. 3, pp. 194–206, 2013.

[46] A. Niculescu, D. Hofs, B. Van Dijk, and A. Nijholt, "How the agent's gender influence users' evaluation of a qa system," in *2010 International Conference on User Science and Engineering (i-USEr)*. IEEE, 2010, pp. 16–20.