# Speaking with virtual humans: Assessing social cognition in traumatic brain injury with a second-person perspective task

Noémie Moreau, Emmanuelle Taché, Maud Champagne-Lavau

# Speaking with virtual humans: Assessing social cognition in traumatic brain injury with a 2nd person-perspective task

Noémie Moreau[1, 2*], Emmanuelle Taché[1*], Maud Champagne-Lavau[1]

1. Aix-Marseille Univ, CNRS, LPL, Aix-en-Provence, France
2. Centre Hospitalier Intercommunal Aix-Pertuis, Service de Neurologie, Avenue des Tamaris, 13616, Aix-en-Provence, France

**Manuscript accepted for publication in *Journal of Neuropsychology***

Moreau, N., Taché, E. & Champagne-Lavau, M. (2021). Speaking with virtual humans: Assessing social cognition in traumatic brain injury with a 2nd person-perspective task. *Journal of Neurospychology*. DOI:10.1111/jnp.12257

**Corresponding author:** Noémie MOREAU
Service neurologie, Centre Hospitalier Intercommunal Aix-Pertuis
Avenue des Tamaris
13616 Aix-en-Provence, cedex 1 France
Phone: (+33) (0)4 42 33 50 71
Fax: (+33) (0)4 42 33 51 58
Email: nmoreau@ch-aix.fr

# Speaking with virtual humans: Assessing social cognition in traumatic brain injury with a 2nd person-perspective task


## Abstract

*Objective*: This study is one of the first to investigate social cognition in participants with traumatic brain injury (TBI) using a task that actively engaged the participant in a real interaction with a partner. Previous results have reported altered social cognition in TBI patients, but social cognition was mostly assessed through traditional tasks involving conscious and deliberate reasoning about characters' mental states (i.e., a third-person perspective). Our goal was to present a new paradigm which allowed the assessment of social cognition in conditions closer to real life meaning that participants were actively engaged in an interaction (i.e., second-person perspective) in order to capture more implicit use of social cognition processes. *Method*: This study used three tasks to evaluate social cognition. We designed a task, called EVICog, in which, participants were engaged in real audiovisual conversations with two virtual humans who expressed emotions and produced speech content that required the participants to make inferences about the characters' mental states. The two other tasks are standard in the literature; they use photographs to test participants' recognition of emotions and short comic strips to test their attribution of intentions. *Results:* Our results showed that TBI participants presented a significant deficit of social cognition compared to control participants. The ROC analysis showed that EVICog has a high discrimination power compared to the other tests. *Conclusion:* These results further confirm that social cognition is altered in TBI participants even in real interactions and further support the use of ecological settings to investigate social cognition.

**Key words**

Social cognition – Traumatic Brain Injury – Real social interactions – Virtual humans – executive functions

## Introduction

Research on social cognition in individuals with traumatic brain injury (TBI) has expanded considerably over recent years and has contributed to a better understanding of patients' relational disorders in daily life. These disorders are particularly detrimental to patients and their relatives and can cause social withdrawal. A number of studies exploring social cognition abilities of individuals with TBI have reported impairments in Theory of Mind (ToM), i.e. the capacity to attribute mental states (e.g., intention, belief, knowledge) to others (Premack & Woodruff, 1978), and in emotions perception (see McDonald, 2013; Radice-Neumann et al., 2007). Despite the central role these abilities play in everyday social interactions, they have traditionally been assessed using tasks in which the participant is required to attribute a mental state (via e.g., false belief understanding, faux pas recognition, non-literal language comprehension such as irony or indirect request) or to perceive an emotion in a character based on short written stories, photos, or comic strips (Fazaeli et al., 2018; Martin-Rodriguez & León-Carrión, 2010). In other words, the participants observed a social interaction and were required to answer questions about the behavior or the emotion of one of the characters but were excluded from taking part in the interaction themselves. These traditional tasks thus do not provide a complete overview of the social cognition abilities of TBI patients as they do not assess how patients actually use their ToM and emotion perception abilities to interact in an appropriate and efficient way with others. The present research focused on the way to improve social cognition assessment in patients with TBI by actively engaging them into a social interaction with a partner.

Researchers have attempted to improve tasks' ecological validity by testing ToM ability and emotion perception with video clips, which are multimodal and dynamic and thus come closer to real-life conditions. Stories, comic strips or photos may be artificial or decontextualized, giving only restrictive information to help people attribute correct mental states or emotions, as opposed to everyday life which offers us a rich context and a larger amount of information (both visual and auditive) (Byom & Mutlu, 2013). In video tasks, TBI patients were shown to be impaired in non-literal language understanding in situations involving irony, sarcasm, deception or lies (Angeleri et al., 2008; Bara et al., 2001; McDonald et al., 2006; McDonald & Flanagan, 2004), or when they were asked to attribute intentions or feelings to actors and judge their attitudes in short movies scenes (Hynes et al., 2011; Turkstra, 2008). Other authors proposed to evaluate ToM abilities in TBI patients interacting with a robot, in a task derived from the Yoni task (Mutlu et al., 2019). In this study, patients had to discover which object among twelve the robot was thinking about by asking him yes-no guessing questions. In some trials, the robot glanced toward the target object, giving a social cue that could be used to attribute the correct mental state to the robot. The results pointed out that the presence of the social cue improved TBI participants' guessing, suggesting that they could use gaze perception as valuable information to attribute mental states. Several studies have also reported deficits in emotion perception with dynamic and/or multimodal material (Drapeau et al., 2017; Marquardt et al., 2001; Zupan & Neumann, 2016). Basic emotions were impacted in The Awareness of Social Inference Test (TASIT; Kelly et al., 2014; McDonald et al., 2006; McDonald & Flanagan, 2004; Williams & Wood, 2010), as well as more subtle emotions such as embarrassment or resentment (May et al., 2017). Compared to traditional tasks, however, the benefit of multimodality is not clear. Some studies report improvements in patients' performances with video tasks (Marquardt et al., 2001; Williams & Wood, 2010; Zupan et al., 2014; Zupan & Neumann, 2016), while McDonald & Saunders, 2005 observed greater difficulties, suggesting that having more information to process (e.g. sentences' semantic content with gestures and visual information) may ultimately hinder correct mentalizing. Overall, except the paradigm of Mutlu et al. (2019), in these tasks, patients were not engaged with other agents but merely observed them in a video and had to answer questions about their mental states, feelings or emotions. While previous studies have attempted to evaluate social cognition in conditions approximating those of everyday life, they have still approached the issue from a perspective which has failed to gauge how ToM functions in real social interactions. Even when social cognition tasks included dynamic material (e.g. videos), participants were merely required to consciously attribute mental states to a fictitious character without being involved in any interaction with him, an approach called "third-person" perspective (de Bruin et al., 2012). For example, Canty et al. (2017) used an interactive virtual environment (i.e., VAMA: Virtual Assessment of Mentalising Ability) where healthy participants were required to visit a virtual shopping center interacting with virtual friends and to complete a list of errands. Their ToM abilities were assessed through multiple-choice explicit questions about their friends' mental states. The multiple-choice included an accurate mentalizing, a reduce mentalizing, an hypermentalizing and a no mentalizing response.

This is however far from the way we use our social cognition in everyday life. Social cognition cannot be summarized as verbalizing others' mental states and perceiving emotions. It sustains our ability to appropriately adapt our behavior and discourse in the course of an interaction with another person in a specific situation in which we are involved in a form of reciprocity. As a consequence, being actively engaged in an interaction with someone, which represents what is known as the "second-person" perspective approach (de Bruin et al., 2012; De Jaegher et al., 2010), may involve different processes compared to when we passively observe an interaction between two people (i.e. "third-person" perspective) (Champagne-Lavau & Moreau, 2013). Third-person perspective social cognition tasks require a type of reasoning which is slow, deliberate, cognitively-demanding and explicit, while we are probably partially aware of our mentalizing activity (Frith, 2012). On the other hand, when we interact with another person (i.e. "second-person" perspective), the way we exert our mentalizing skills is most likely implicit, fast and automatic. For example, we receive feedback from our interlocutors which may impact on how we attribute mental states and the dynamics of our interaction (de Bruin et al., 2012; De Jaegher et al., 2010; Schilbach et al., 2013). Considering this distinction between a third-person perspective and a second-person perspective approaches of social cognition, it appears that the assessment of social cognition, i.e., a function that is essential to maintaining appropriate social relationships and adapting our behavior to others, in TBI patients needs to be improved using a “second-person” perspective task which actively engages them in real interactions.

With this goal in mind, we developed an original paradigm called EVICog (in French: *Evaluation de la Cognition sociale en interaction Virtuelle*) in which the participant takes part in a real audio-visual conversation about everyday situations with virtual humans. Using virtual humans makes it possible to evaluate social cognition, in real social interactions while retaining good experimental control, which is difficult to achieve in interactions with real humans (Bente et al., 2001; Blascovich et al., 2002; Byom & Mutlu, 2013; Canty et al., 2017; Schilbach et al., 2013). Furthermore, the participant is given the socio-emotional impression of having a real conversation with someone else (Bente et al., 2001). Studying real interactions is particularly important for TBI individuals as they are frequently reported to have social relationship disorders (Engberg & Teasdale, 2004; Hawthorne et al., 2009; Stone & Hynes, 2011; Temkin et al., 2009). These disorders are among the symptoms which have the most deleterious impact on the quality of life of patients and their relatives (Dahlberg et al., 2006; Kinsella et al., 1991; Koskinen, 1998).

## Objective

The main purpose of this study was to present a novel second-perspective task allowing the assessment of social cognition abilities (i.e. Theory of Mind, emotion perception) in participants with TBI, participants being actively engaged in a real conversation with a virtual partner. The EVICog task enabled us to test - in French - how participants implicitly attribute mental states and perceive emotions in the course of a reciprocal interaction by analyzing their replies to the virtual human's utterances. We also used traditional third-person perspective tasks in which the participant have to attribute mental states and emotions without being involved in any interaction, in order to examine the relationship between our new task and the standard ones.

We hypothesized that participants with TBI would show social cognition difficulties when they were tested using the EVICog task and the third-person perspective tasks. We also expected that patients' performances on the EVICog task would significantly correlate with scores of standard social cognition tasks.

## Method

### Participants

Twenty eight individuals with TBI (22 men, 6 women; age *M* = 37.4; *SD* = 9.7 years, educational level *M* = 12.4; *SD* = 2 years) participated in our study. They had sustained a moderate to severe traumatic brain injury (1 moderate and 27 severe TBI). They were compared to 31 healthy control participants (HC) (22 men, 9 women; age *M* = 37.6; *SD* = 11.2 years, educational level *M* = 13.1; *SD* = 2.9 years). TBI participants and HC participants were matched for age ($t(57) = 0.026$, $p > .05$) and educational level ($t(57) = - 1.13$, $p > .05$). All participants were between 18 and 60 years old, native French speakers, and had no history of psychiatric disorders. The TBI participants were recruited by way of various regional organizations which serve the TBI community in terms of evaluation, rehabilitation and professional orientation (SAMSAH TC-CL, UEROS Centre Phocée), or through a local association for individuals with TBI (Groupe d'Entraide Mutuelle). Due to our recruitment procedure, access to imagery data was not available. Our inclusion criteria were: 1) The brain injury occurred after the age of 18, 2) at least one year had passed between the TBI and participation in the study. The exclusion criteria for our study were: 1) aphasia, 2) visual or hearing disorders, 3) chronic alcoholism or addictions, 4) a history of other neurological disorders or several TBIs. The severity of the TBI was determined using the criteria developed by Teasdale & Jennett (1974): mild TBI for a Glasgow scale score greater than 13; moderate TBI for a score between 9 and 12 and severe TBI when the score was less than 9. The Glasgow score was available for 17 TBI participants (*M* = 5.5, *SD* = 2). When this score was absent, the severity of the TBI was evaluated by clinicians from clinical data such as coma duration (*M* = 74, *SD* = 86 days). Mean time since injury was 122 months (*SD* = 84). Injuries were caused by motor vehicle accidents (n= 26), sporting accidents (n = 1), and assault (n = 1). All participants gave their written informed consent before recruitment. The study was approved by the regional independent ethics committee (CPP Sud-Méditerranée I) and by the French Agency for Health Product Safety (ANSM).

## Materials

### *Neuropsychological assessment*

All participants were evaluated on memory (i.e., autobiographical memory, destination memory, semantic memory) and executive functioning (i.e., inhibition, shifting, access, updating). The participants' autobiographical memory was assessed with a task testing autobiographical fluency (AF) (Piolino et al., 2010) in which participants were asked to list as many autobiographical memories as they could according to four levels of specificity (two minutes were given for each of the following levels): (AF1) events during their lifetime which lasted for at least three years (e.g. my relationship with Mark, my studies in Paris, etc…), (AF2) general events within one previously cited period of time lasting for a few days or weeks (e.g. my trip in Italy with Mark), (AF3) specific and one-time events which occurred during one previously cited general event and lasted for a few hours or a day (e.g. this fantastic day in Rome), (AF4) specific details (perceptions, emotions, etc…) which occurred during one previously cited specific event (e.g. the visit of the Vatican, the crowd in the Sistine Chapel, the beautiful paintings on the ceiling…). The total number of cited events in the four conditions was used as an autobiographical memory index. Destination memory was assessed using a paradigm designed by El Haj et al. (2015) in which participants were asked to read proverbs associated with different celebrities out loud (one proverb for each celebrity) in the learning phase. In the recognition phase, proverb-celebrity pairs were presented to participants who had to decide whether they had previously assigned that proverb to that celebrity or not. The total number of correct answers was used as a destination memory index (max = 24). The Information Subtest of the Wechsler Adult Intelligence Scale (Weschler, 2011), with the total number of correct answers, was used to assess semantic memory (max = 26). Executive functioning was evaluated by: 1) the Hayling Completion Test, part A (automatic) and part B (inhibition) (Burgess & Shallice, 1997) leading to two scores (part A and part B) calculated according to the method developed by Burgess and Shallice (1997). Response latencies for each part and response accuracy in part B were taken into consideration in these calculations; 2) the Trail Making Test (Godefroy et al., 2008) with the difference between the completion of part B and part A as a shifting index; 3) the semantic, phonemic and alternated fluency of the Delis-Kaplan Executive Function System (DKEFS, Delis et al., 2001) with the total number of words produced in each condition as access indexes; and 4) the three digit spans of the Wechsler Adult Intelligence Scale (Wechsler, 2011): forward span, backward span and ordered span. The total number of correct sequences repeated for each span was considered as updating index.

***Social cognition assessment***

*Third-person perspective assessment*

The participants' social cognition abilities were assessed with traditional third-person perspective tasks and a novel second-person perspective task (EVICog). In the third-person perspective, ToM ability was assessed using the Character Intention Task (CIT) (Sarfati et al., 1997). Participants were asked to choose from three pictures, one of which appropriately completed a comic strip according to the character's mental state. The dependant variable was the number of correct responses (max = 28). We also evaluated emotion recognition using the Ekman Faces test (EFT). Forty two pictures of the following 6 basic emotions and neutral expression were selected from the Ekman & Friesen (1976) collection: anger, sadness, fear, disgust, surprise, joy, neutral. Participants were asked to choose the emotion that best described the picture using a force-choice paradigm. We measured the number of correct responses (max = 42).

*Second-person perspective assessment*

For the second-person perspective assessment, an original ecological task called EVICog was designed for the purpose of this study. This task allowed us to assess both ToM ability (i.e., the attribution of knowledge/desires, the attribution of ironic intent) and emotion perception in a social interaction between the participant and a virtual character. In this task, the participant was engaged in nine conversations with two virtual humans, called Pierre and Marie, one at a time. The software simulated a video call on the computer and looked like well-known video call applications. The topics of the conversations were everyday subjects such as which birthday gift to buy for a friend or a trip to the movies. Each conversation followed a pre-programmed script in which each of the virtual human's replies was predetermined by means of the Wizard of Oz technique[1].

*Stimuli*. Nine scenarios were created, five with Marie and four with Pierre. Three types of items were used to assess one of the domains of social cognition: emotion perception or Theory of Mind abilities (i.e., the attribution of ironic intent to the speaker and the attribution of knowledge/desires). There were six items with emotional content, six items designed to test the attribution of ironic intent and six items to test the attribution of knowledge /desires for a total of 18 stimuli (see appendix for examples of each item and transcripts of conversations between a control participant or a TBI participant with the virtual agent). Each scenario contained two target sentences spoken by one of the virtual humans, while the other sentences uttered by the virtual humans were fillers leading to more natural and fluid conversations. The target sentences were the same for each participant, whereas the fillers could vary from one participant to another depending on what the participant said during the conversation. Only the participant's replies to the target sentences were analyzed. For the emotion perception items, the virtual human expressed an emotion (via prosody and facial expressions) while he/she was talking to the participant. The participant was required to perceive the emotion and reply to the virtual human in an appropriate way (see appendix for examples). Six emotions were used in this evaluation (i.e., disgust, anger, surprise, sadness, anxiety, and disappointment). These emotions were chosen according to the topics of the conversations. In the items measuring the attribution of ironic intent to the speaker, the virtual human said an ironic utterance during the conversation and the participant was required to attribute the right mental state to the virtual human (ironic intent) and thus understand that the virtual human did not really mean what he/she said. The attribution of ironic intent depended on whether the

[1] This technique allows an individual to interact with a computer system (e.g., the virtual human in our experiment) without knowing that responses given by the system are simulated by a human. Individuals interacting with the computer system believe that this system is autonomous (Kelley, 1984).

participant picked up on an incongruity between the target sentence and different information mentioned during the conversation, as well as his/her perception of ironic prosody. Correct attributions of ironic intent to the virtual human led the participant to reply appropriately (see appendix for an example). In the items measuring the attribution of knowledge/desires, the participant was required to use his/her knowledge about the situation and attribute someone else's desires or knowledge (either encyclopedic knowledge or knowledge about the interlocutor's preferences) to appropriately answer specific questions asked by the virtual human during the conversation (see appendix for an example). The different conversational scripts were presented in the same order for each participant, since the conversations were connected to each other (e.g., Pierre gives important information in one conversation which is useful to correctly attribute mental states in following conversations).

*Software*. The virtual humans, Pierre and Marie, were visually designed for this study using Mixamo® software. Their facial movements and expressions were captured from a real human face with a Primesense Carmine® 3D camera and were superimposed directly onto the virtual humans with Faceshift® software. Our virtual humans had very fluid head movements, realistic facial expressions and real human voices synchronized with lip movements. Their voices were real human voices - every possible sentence that could be said by the virtual humans during the conversations was audio recorded by a male and female volunteer. These volunteers were asked to produce the sentences following instructions about the prosody (e.g. using an ironic tone, emotion, anger, sadness, etc).

*Procedure*. The EVICog task lasted about half an hour and was administered in a quiet room. The participant sat in front of a screen (connected to computer of the experimenter who controlled the conversational flow) to receive the video calls while the experimenter monitored the software on a laptop. The participant could not see the experimenter's screen or the fact that the experimenter was controlling the interaction and choosing the virtual human's replies. The conversations were audio-recorded using a Zoom H4 recorder, and the participant wore a headset microphone. This gave the participant the impression that their interlocutor could hear their replies. The participant was given the following instructions: “You are going to talk to a friend of yours, called Marie or Pierre (one at a time), via a Skype-like video call. Your conversations will be short and focus on subjects from daily life. Imagine that Pierre and Marie are old friends of yours, and it is very important for you to talk to them very naturally, as you would with your own friends. In some cases, Pierre or Marie will start the call; in other cases, you will. Before each call, you will be told what the topic of the conversation will be, as well as the identity of your interlocutor (Pierre or Marie) and who is going to start the call (you or the interlocutor)”. Before each of the nine conversations, the experimenter explained to the participant who would call whom (for five of the conversations, Pierre or Marie called the participant), what general topic would be addressed (ex: “You have plans with Pierre this afternoon; he is going to call you to ask you what you want to do and you will answer that you would like to go to the movies”). The virtual human ended every conversation by saying that he/she had to go and that they would talk again later. Each conversation lasted less than two minutes. A practice conversation was administered to ensure that the participant understood the task and was familiar with the material. The conversations followed a preprogrammed script: during the conversation, the experimenter clicked on the best reply the virtual human could say, among several possibilities, depending on what the participant had just said (Cf. Figure 1). The lists of the virtual humans’ replies were as exhaustive as possible, and they were compiled with data from a pilot study[2] with healthy volunteers. When the

[2] A pilot study was conducted with 37 healthy participants who were native French speakers (mean age: 26.4, range: 18-56; mean educational level: 14.2, range: 12-20). This study allowed us to ensure that more than 70% of the participants were able to give appropriate replies to the target sentences during the nine conversations, to improve fluidity of conversations and enlarge the number of pre-registered fillers to cover a maximum range of possible replies.

experimenter clicked on the response, the virtual human said the corresponding sentence. The participant was unaware that the experimenter was involved throughout the experiment, and they reported believing that the virtual human was able to reply “on his/her own”, in a post-experiment interview.

*Figure 1 around here*

To ensure that participants who failed to reply correctly to the attribution of knowledge /desires task had the required encyclopedic knowledge (e.g., see the scenario “*Invitation for lunch*” in the appendix) and to control for memory effects, control questions were asked to the participants at the end of the EVICog task. These questions focused on encyclopedic knowledge (e.g., Is couscous a Mexican dish?; scenario “*Invitation for lunch*” in the appendix) or information previously mentioned by one of the virtual humans during the task (e.g., Does Pierre like going to the movies?; scenario “*Gift for Pierre’s birthday*” in the appendix).

*Scoring*. Speech productions were transcribed verbatim. Participants’ productions were scored by two raters, one of them independent from the study. Only participants’ replies to target sentences were analyzed. Their replies to the target utterances were scored 1 if they appropriately reflected a clear understanding of the situation (meaning that they perceived the emotion, correctly attributed ironic intent, or correctly attributed knowledge or desires) and 0 when the participant’s reaction was not appropriate (see examples of scoring in appendix). More precisely, for the emotion perception domain, participants’ replies were scored 1 if they referred to the emotion expressed, either directly (e.g., “you don’t want to, do you?”), or indirectly (e.g., “what happened, you don’t want to go after all?”). Other replies that did not take into account the emotion expressed (e.g., “Yes”) were scored 0. For the attribution of ironic intent domain, participants’ replies were scored 1 when participants correctly interpreted sentence of the virtual human as meaning the opposite or something different from what he/she said (e.g., "hum, you don't seem in the mood") or as mocking someone or him/herself. Other replies reflecting a literal interpretation of the virtual human sentence (e.g., "great, what movie do you want to see?") were scored 0. For the attribution of knowledge/desires domain, the virtual human said something inaccurate about general knowledge (e.g., Pierre saying about couscous: "I'm sorry but I'm not fond of Mexican food") or about the other virtual human desires (e.g., gift for a birthday). Participants' replies were scored 1 when participants corrected the virtual human error. Other replies that did not reflect a correct attribution of knowledge or desire to the virtual human (e.g., when participants agreed with the virtual human proposition) was scored 0. Cohen's Kappa coefficients were calculated to estimate inter-rater agreement for the 18 target sentences analyzed (6 targets x 3 domains) for all the participants. The inter-rater reliability was 94.1% with Cohen's Kappa $k = .861$, $p < .0001$. The maximum score was 6 for each social domain (i.e., emotion perception and Theory of Mind abilities: attribution of ironic intent to the speaker, attribution of knowledge/desires). Accuracy for the attribution of knowledge/desires domain was analyzed only for those items for which the post EVICog task control questions were answered correctly. Thus, for each item, data from participants who incorrectly answered the control question was removed from the analyses. As a consequence, scores were converted to percentages for the statistical analyses.

**Data analyses**

Unpaired t-tests were used to explore group differences on the neuropsychological variables (executive functions and memory), on the total number of correct responses on the CIT and on the percentages of correct responses in the three domains tested by the EVICog task (i.e., emotion perception, attribution of ironic intent, attribution of knowledge/desires). A 2-group (TBI, HC) x 4-condition (AF1, AF2, AF3, AF4) repeated-measures ANOVA was performed on the scores of the autobiographical fluency task. To determine the differences between groups in the Ekman Faces test, a 2-group (TBI, HC) x 7-emotion (anger, disgust, sadness, fear, joy, surprise and a neutral emotion) repeated-measures ANOVA was performed on the total number of correct responses. Spearman correlation analyses were performed to examine the relationship between the three domains of EVICog and the corresponding standard measures of social cognition. A ROC (Receiver Operating Characteristics) curve analysis was also done for the different measures of social cognition to test for classification accuracy of the EVICog task against the standard tests. The accuracy of the ROC curve was quantified by the area under the curve (AUC). An AUC of 0.50 means that a test's diagnostic performance is equal to chance, while an AUC of 1.0 means perfect diagnostic performance. Measures of effect size were calculated for each effect of interest by providing the partial eta-squared for ANOVAs and the Cohen's d for t-test. The alpha level was set at $p < .05$ for all the analyses.

## Results

### Group comparison on neuropsychological measures

The neuropsychological characteristics of both groups are presented in Table 1. Participants with TBI showed lower performances than HC participants on all executive function measures, showing deficits in inhibition, shifting, access and working memory. Participants with TBI also obtained lower scores in destination memory and in semantic memory as measured by the Information subtest. Finally, the results of the 2-group × 4-condition (levels of specificity of the memories) repeated-measures ANOVA on the total number of events produced in the autobiographical fluency task showed a main effect of condition ($F$ (3, 168) = 51.42, $p$ <.0001, $\eta^2$=0.48). Participants produced more events in AF4 compared to AF1 ($p$ <.0001), AF2 ($p$ <.07) and AF3 ($p$ <.0001), more events in AF2 compared to AF1 ($p$ <.02) and AF3 ($p$ <.0001) and more events in AF1 compared to AF3 ($p$ <.0001). There was also a main effect of group (*F* (1, 56) = 97.29, $p$ <.0001, $\eta^2 = 0.63$) showing that HC participants gave more autobiographical events compared to participants with TBI. The interaction group x trial was significant (*F* (3, 168) = 8.34, $p$ <.0001, $\eta^2 = 0.13$) and was decomposed by group. In the HC group, the pattern was similar to that previously described; more events were produced in AF4 than in AF1 ($p$ <.0001), AF2 ($p$ <.01) and AF3 ($p$ <.0001); more events were produced in AF2 than in AF1 ($p$ <.01) and AF3 ($p$ <.0001); and more events were produced in AF1 than in AF3 ($p$ <.0001). In the TBI group, there was no difference ($p$ >.05) between AF1, AF2 and AF4, but TBI participants produced fewer events in AF3 ($p$ <.0001).

*Table 1 around here*

## Group comparisons on social cognition

### *Third-person perspective tasks: the CIT and Ekman Faces test*

In the CIT, participants with TBI obtained lower performances compared to HC participants ($t$(57) = -4.595, $p$ <.0001, Cohen's $d$ = 1.32) (TBI participants: $M$ = 24, $SD$ = 4.36; HC participants: $M$ = 27.87, $SD$ = 0.43). The 2 groups x 7 emotions repeated-measures ANOVA performed on the number of correct responses in the Ekman Faces test revealed a main effect of emotion type (*F* (6, 336) = 67.517, $p$ < .0001, $\eta^2$ = .55). Joy was recognized better than all the other emotions including the neutral emotion ($p$ <.0001); surprise and the neutral emotion (no difference, $p$ >.05) were recognized better than the four negative emotions ($p$ <.0001); anger and disgust (no difference, $p$ >.05) were recognized better than sadness and fear ($p$ <.0001); and finally, sadness and fear (no difference, $p$ >.05) were recognized less often than all the other emotions and the neutral emotion ($p$ <.0001). There was also a main effect of group (*F* (1, 56) = 7.09, $p$ < .01, $\eta^2$=0.112) showing that participants with TBI performed worst than HC participants. The interaction group x emotion was not significant ($F$ (6, 336) = 0.302, $p$ >.05). Thus, participants with TBI obtained lower scores on all the emotion recognition tasks compared to control participants, but the pattern of correct responses among the different emotions was similar in both groups (Cf. Figure 2).

*Figure 2 around here*

***Second-person perspective task: EVICog***

Unpaired t-tests performed on the percentage of appropriate replies in the three domains tested by EVICog showed that HC participants outperformed participants with TBI in emotion perception ($t(57) = -5.426$, $p$ <.0001, Cohen's $d$ = 1.47), attribution of ironic intent to the speaker ($t(57) = -4.909$, $p$ <.0001, Cohen's $d$ = 1.35), and attribution of knowledge/desires ($t(57) = -6.160$, $p$ <.0001, Cohen's $d$ = 1.67) (Cf. Figure 3). A ToM score from EVICog was calculated by averaging percentages of appropriate responses from the attribution of ironic intent domain and the attribution of knowledge/desires domain. Groups comparison on this ToM score also showed better performances in the HC participants than participants with TBI ($t(57) = -7.176$, $p$ <.0001, Cohen's $d$ = 1.95).

*Figure 3 around here*

***Correlation between second and third-person perspective tasks***

Results of the Spearman correlation analyses performed in the TBI group, showed that there was a significant correlation between the Ekman Faces test score and the EVICog emotion score (r = .528, p < .002), between the CIT score and the attribution of ironic intent score (r = .676, p <. 0001), between the CIT score and the attribution of knowledge/desires score (r = .340, p< .041) and between the CIT score and the ToM score (r = .674, p<.0001).

***Receiver Operating Characteristic (ROC) curve***

The AUC for the social cognition measures were 0.902 for the ToM domain of EVICog (including attribution of ironic intent, attribution of knowledge/desires), 0.842 for the emotion perception domain of EVICog, 0.849 for the CIT and 0.644 for the Ekman faces test (see figure 4). These results reflect that the sensitivity and specificity was very high for the ToM domain of EVICog and high for the emotion perception domain compared to the CIT and the Ekman Faces test.

*Figure 4 around here*

**Correlation between second second and third-person perspective tasks and neuropsychological measures**

Spearman correlation analyses were conducted in the TBI group between score of each domain tested by EVICog (i.e., emotion perception, ToM), the CIT, the Ekman faces test and the neuropsychological measures (i.e., executive functions, memory). Overall, the results showed that autobiographical fluency total score, semantic memory score and destination memory were associated with performances in emotion perception and ToM in both EVICog and classical tasks. However, there were some differences between the EVICog task and the classical tasks. A relationship was found between emotion perception in EVICog and flexibility, short term memory and access (phonemic fluency) while this was not the case for the Ekman faces test. Regarding ToM abilities, inhibition was shown to be associated with the CIT but not with the EVICog task (Cf. Table 2 for details of the results).

*Table 2 around here*

To sum up, participants with TBI were impaired on both traditional tasks evaluating social cognition in a third-person perspective and on the EVICog task evaluating social cognition (i.e., emotion perception, attribution of ironic intent, attribution of knowledge/desire) in a second-person perspective. All domains of the EVICog task demonstrated adequate correlations with standard third-person perspective tasks of social cognition in participants with TBI. ROC curves indicated that the EVICog task has an adequate discriminative capacity for detecting social cognition impairments in participants with TBI.

## Discussion

This study investigated social cognition in participants with TBI using a new interactive task, called EVICog, which actively engaged participants in conversations with virtual humans in order to evaluate social cognition deficits in TBI individuals in conditions resembling everyday interactions. Our main results showed that EVICog is a sensitive tool to detect social cognition impairments in participants with TBI when they were actively engaged in a real interaction. Participants with TBI were found to be particularly impaired in real social interactions when they were asked to attribute an emotion, an ironic intention, knowledge or desires to virtual characters and behave accordingly to maintain an efficient interaction. Our study also confirmed previous research showing ToM impairment (Fazaeli et al., 2018; Martin-Rodriguez & Leon-Carrion, 2010) and emotion perception impairment (McDonald, 2013; Radice-Neumann et al., 2007) in TBI individuals in traditional third-person perspective tasks (e.g., the Ekman Faces test and the CIT).

All social cognition tests separated participants with TBI from healthy participants. Overall, the observation of correlations between the different scores of EVICog and other standard measures of social cognition showed that the EVICog task has a good validity to evaluate social cognition. In addition, comparisons of the areas under the ROC curves pointed out that EVICog was a better test in discriminating the groups. Thus, the EVICog task is a valuable tool for assessing social cognition impairments. It gives the advantage to measure social cognition in a more naturalistic way, closer to real-life interactions without asking for verbal explanations of other's mental states. The present study was designed to explore social cognition in conditions where participants were actively involved in real interactions engaging reciprocity and feedbacks between them and their virtual partners. They spoke naturally with the virtual characters as if they were friends of theirs, which made it possible to access an implicit form of social cognition, or a type of processing similar to that used in everyday social interactions. Contrary to the CIT or the Ekman Faces test, EVICog did not require any explicit representation or verbal explanation of another person's behavior, and thus could give access to a use of social cognition closer to real-life conditions compared to the tasks employed in third-person perspective (Champagne-Lavau & Moreau, 2013; de Bruin et al., 2012; Schilbach et al., 2013). In the EVICog task, the participants did not know that we were interested in their ability to attribute an emotion or other mental states to the virtual humans and react appropriately in the conversations; they were only instructed to speak naturally to their interlocutor as they would in usual, everyday conversations. As a consequence, they were required to answer their interlocutor quickly and react accordingly to their feedbacks throughout the conversation (Apperly et al., 2010; Byom & Mutlu, 2013; Schilbach et al., 2013). The participants also needed to take into account relevant environmental cues necessary to interact efficiently and behave appropriately, as described by Achim et al. (2013). On third-person perspective tasks, participants could use much more time to observe the stimuli, analyze the situation, prepare and choose their answers, which may not reflect how social cognition functions during real social interactions (Byom & Mutlu, 2013). Because these tasks offer extra time to respond, they may overestimate TBI participants' ToM performances compared to their deleterious functioning in daily life. Our results thus highlight the fact that immersion in real interactions and mentalizing in real life may involve implicit and spontaneous ToM processes since it engages reciprocity and feedbacks between partners, contrary to third-person perceptive tasks (including attributing mental states to characters in videos).

Thus, EVICog seems to be a relevant tool to evaluate social cognition and the impairments that TBI individuals face in terms of their social abilities. It offers promising perspectives about the use of virtual technology in neuropsychological evaluation and maybe rehabilitation, since these "ecological" new tools can immerse individuals in a large range of situations resembling real life. Moreover, in a post-experiment interview, a large majority of our control and TBI participants reported feeling totally satisfied with their interactions and were under the impression of having a real conversation with the virtual humans.

Some limitations have nevertheless to be mentioned. The EVICog paradigm did not allow us to disentangle between the patients' ability to correctly perceive emotions and their ability to appropriately react to the virtual humans' emotions. Even if the literature showed poor facial and vocal emotion recognition in patients with TBI (Babbage et al., 2011; Zupan et al., 2014), further research is required to clarify this issue. Although the procedure used to create the virtual humans allows fluid head movements, realistic facial expressions and real human voices synchronized with lip movements (see software section), virtual humans are not as sophisticated as real humans and their facial expressions may be perceived as less detailed and realistic than in real people which could have impacted the way participants perceived and reacted to virtual humans' emotions.

## Conclusion

The present research has provided the first experimental evidence that social cognition is impaired in participants with TBI when they were actively engaged in real interactions such as conversations with a partner. Our new paradigm is of interest since it simulates the way social cognition is used in social interactions and thus can help us to better understand TBI individuals' social behavior impairments in everyday life. The EVICog task was designed to identify TBI individuals' difficulties perceiving emotions, understanding irony and attributing mental states to others and, as a consequence, their struggles to react appropriately to others' feedback and interact appropriately with others in social settings. This study may also open perspectives about the use of virtual technologies as a useful tool to support cognitive remediation and social rehabilitation.

# Appendix

## Examples of stimuli from the EVICog task

### *Example of a stimulus used to perceive emotions:*

The participant asks Pierre if he wants to go to the movies and Pierre answers: “Yeah, if you want to,” but with an expression of disgust.

In this example, the participant was intended to understand that Pierre does not really want to go to the movies, and his/her answer needed to show that he/she understood Pierre’s real emotional state (ex: “You don’t want to, do you ?”).

### *Example of a stimulus used to attribute ironic intent:*

About going to the movies, Pierre says: “Paying to go watch a movie in an uncomfortable seat with people around speaking and making noise! It’s great!” with an ironic prosody.

In this example, the participant was expected to understand that Pierre says the opposite of what he really means.

### *Example of a stimulus used to attribute desires/knowledge:*

The participant invites Pierre to have lunch together and proposes to prepare a couscous. Pierre answers “I’m sorry but I’m not fond of Mexican food”.

In this example, the participant was expected to correct Pierre’s mistake about the origin of couscous.

**Transcripts** (in French and in English) of conversations between a participant and the virtual agent in the EVICog task (the replies in italics were analyzed):

1. ***Scenario about "Going to the movies" with an expression of emotion (disgust) and ironic intent. Conversation between a control participant and the virtual agent.***

Pierre vous appelle car vous devez vous voir cet après-midi mais vous n'avez pas encore décidé de ce que vous allez faire. Il va vous demander ce que vous avez envie de faire et vous répondrez que vous aimeriez aller au cinéma.

- Pierre: Salut !
- Participant : Salut Pierre !
- Pierre: Alors qu'est-ce qu'on fait cet après-midi ?
- Participant: Et ben on peut aller au cinéma ?
- Pierre: Ouais si tu veux (the face of the virtual agent shows an emotion of disgust).
- Participant: *T'as pas trop envie ?* (➔ score = 1)
- Pierre: Non pourquoi pas, moi j'adore aller au cinéma. Payer pour regarder un film dans un siège pas confortable, avec les gens autour qui parlent et font du bruit. C'est génial ! (Irony).
- Participant: *Mmh, t'es pas trop chaud.* (➔ score = 1)
- Pierre: Oui voilà, j'aime pas aller au cinéma, je préfère qu'on fasse autre chose. Mais c'est pas grave on verra cet après-midi ce qu'on a envie de faire. Allez faut que j'y aille ! A tout à l'heure !
- Participant: Salut !

Pierre is calling you because you decided to see each other this afternoon. He is going to ask you what do you want to do, and you are going to answer that you feel like going to the movies.

- Pierre: Hi!

- Participant: Hi, Pierre !

- Pierre: So, what are we going to do this afternoon?

- Participant: Well, we could go to the movies?

- Pierre: Yeah, if you want to (the face of the virtual agent shows an emotion of disgust).

- Participant: *You don't want to, do you?* (➔ score = 1)

- Pierre: (replying ironically): Oh, you know, I love going to the movies. Paying to go watch a movie in an uncomfortable seat with people around speaking and making noise! It's great! (Irony)

- Participant: *Huh, it sounds like you don't feel like it*. (➔ score = 1)

- Pierre: Yeah, that's it, I don't like going to the movies. I would rather do something else. We'll see this afternoon. I need to go. See you later!

- Participant: Bye!

2. ***Scenario about a "gift for Pierre's birthday" with an expression of emotion (surprise) and attribution of knowledge/desire. This scenario follows (not immediately after) a conversation between Pierre and the participant in which Pierre explained that he is not really fond of going to the movies. Conversation between a TBI participant and the virtual agent.***

Marie vous appelle car c'est bientôt l'anniversaire de Pierre et vous avez décidé de lui acheter un cadeau ensemble. Elle va commencer la conversation en vous demandant si vous avez des idées de cadeaux. Vous lui répondrez que vous avez pensé à une montre.

- Marie: salut !

- Participant: bonjour.

- Marie: Alors dis-moi, tu as eu une idée pour le cadeau d’anniversaire de Pierre ?

- Participant: oui une montre.

- Marie: une montre ??? (with an emotion of surprise).

- Participant: *oui. Ce qui donne l’heure* (➔ score = 0, no reaction to surprise)

- Marie: Pourquoi pas. C’est une bonne idée mais il me semble qu’il m’avait dit qu’il n’aimait pas tellement les montres, et d’ailleurs je le vois pas souvent en porter une. Sinon moi j’avais pensé lui prendre un abonnement à l’année au cinéma. Parce que je sais qu’il adore regarder des films donc comme ça il pourra les voir dès leur sortie. T’en penses quoi, tu penses que ça plaira à Pierre ? (Marie ignores that Pierre doesn’t like going to the movies).

- Participant: *non c’est pas intéressant*. (➔ No scoring, need further precision)

- Marie: ah bon pourquoi tu dis ça ? (asking for precision)

- Participant: *parce qu’il y a des films toute l’année*. (➔ score = 0, no attribution of knowledge to Marie, i.e. the fact that she ignores Pierre doesn’t like going to the movies).

- Marie: ah d’accord bon on laisse tomber cette idée alors. Par contre j’avais pas vu l’heure, il faut que je file mais on en reparle plus tard, à bientôt.

- Participant: au revoir à bientôt !

Marie calls you because Pierre’s birthday is approaching and you have decided to offer him a gift together. She will begin the conversation asking you if you have some ideas for a gift. You will answer her that you have thought about a watch.

- Marie: Hi!

- Participant: Hello.

- Marie: So tell me, do you have any idea for Pierre’s birthday gift?

- Participant: yes, a watch.

- Marie: a watch??? (with an emotion of surprise).

- Participant: *yes. The thing that gives the hour* (➔ score = 0, no appropriate reaction to surprise)

- Marie: why not? It is a good idea but I think that he told me once that he didn't like watches, and I rarely see him wearing a watch by the way. I rather thought that we could offer him a season ticket for the cinema for one year. Because I know he loves seeing movies so he will be able to watch them as soon as they will be released. What do you think about that ? Do you think Pierre will enjoy that? (Marie ignores that Pierre doesn't like going to the movies).

- Participant: *no, it is not interesting.* (➔ No scoring, need further precision)

- Marie: really? Why do you say that? (Asking for precision)

- Participant: *because there are movies all year round* (➔ score = 0, no attribution of knowledge to Marie, i.e. the fact that she ignores Pierre doesn't like going to the movies).

- Marie: ok, let's forget this idea. Oh, I didn't see the hour, I need to go but we will discuss this later. See you soon!

- Participant: Good bye, see you later!

3. ***Scenario about an "invitation for lunch" with the expression of emotion and attribution of knowledge/desire. Conversation between a TBI participant and the virtual agent.***

Vous appelez Pierre pour lui proposer de venir manger chez vous ce midi. S'il vous demande quel plat vous avez prévu de faire, vous lui répondrez du couscous.

- Pierre: Salut !

- Participant: Bonjour. Tu viens manger à la maison ce midi ?

- Pierre: Ouais bien sûr, super ! qu'est-ce que t'as prévu de bon à manger ?

- Participant: Du couscous.

- Pierre: Ah du couscous (with an emotion of disappointment).

- Participant: *Et oui* (➔ score = 0, no reaction to disappointment)

- Pierre: Je suis désolé mais en fait j'aime pas du tout la cuisine mexicaine, c'est trop épicé pour moi. (false knowledge about the origin of couscous).

- Participant: *C'est pas mexicain, c'est maghrébin.* (➔ score = 1)

- Pierre: Ah mais oui, j'ai confondu avec autre chose en fait, alors c'est bon je peux en manger. Bon faut que je te laisse, mais on se voit à midi alors ! A tout à l'heure !

- Participant: A tout à l'heure !

You are calling Pierre to have lunch with him. If he asks what you're having, you'll answer that you're making couscous.

- Pierre: Hi!

- Participant: Hello! Do you want to come over for lunch today?

- Pierre: Sure, of course, great! What are you making?

- Participant: Couscous.

- Pierre: Oh… couscous (with an emotion of disappointment).

- Participant: *Yes* (➔ score = 0, no reaction to disappointment)

- Pierre: I'm sorry but I'm not fond of Mexican food. It's too spicy for me (false knowledge about the origin of couscous)

- Participant: *It's not Mexican, it's North African* (➔ score = 1)

- Pierre: Oh yeah, you're right! I confused it with something else. It's fine, I can eat that! Well, I need to go. I'll see you at noon!

- Participant: See you later!

# References

Achim, A. M., Guitton, M., Jackson, P. L., Boutin, A., & Monetta, L. (2013). On What Ground Do We Mentalize? Characteristics of Current Tasks and Sources of Information That Contribute to Mentalizing Judgments. *Psychological Assessment*. https://doi.org/10.1037/a0029137

Angeleri, R., Bosco, F., Zettin, M., Sacco, K., Colle, L., & Bara, B. (2008). Communicative impairment in traumatic brain injury: A complete pragmatic assessment. *Brain and Language*, *107*(3), 229–245. https://doi.org/10.1016/j.bandl.2008.01.002

Apperly, I. A., Carroll, D. J., Samson, D., Humphreys, G. W., Qureshi, A., & Moffitt, G. (2010). Why are there limits on theory of mind use? Evidence from adults' ability to follow instructions from an ignorant speaker. *The Quarterly Journal of Experimental Psychology*, *63*(6), 1201–1217. https://doi.org/10.1080/17470210903281582

Babbage, D. R., Yim, J., Zupan, B., Neumann, D., Tomita, M. R., & Willer, B. (2011). Meta-analysis of facial affect recognition difficulties after traumatic brain injury. *Neuropsychology*, *25*(3), 277–285. https://doi.org/10.1037/a0021908

Bara, B. G., Cutica, I., & Tirassa, M. (2001). Neuropragmatics: Extralinguistic Communication after Closed Head Injury. *Brain and Language*, *77*(1), 72–94. https://doi.org/10.1006/brln.2000.2430

Bente, G., Krämer, N. C., Petersen, A., & De Ruiter, J. P. (2001). Computer Animated Movement and Person Perception: Methodological Advances in Nonverbal Behavior Research. *Journal of Nonverbal Behavior*, *23*(5), 151–166.

Blascovich, J., Loomis, J., Beall, A. C., Swinth, K. R., Hoyt, C. L., & Bailenson, J. N. (2002). TARGET ARTICLE: Immersive Virtual Environment Technology as a Methodological Tool for Social Psychology. *Psychological Inquiry*, *13*(2), 103–124. https://doi.org/10.1207/S15327965PLI1302_01

Burgess, P. W., & Shallice, T. (1997). *The Hayling and Brixton Tests. Test manual* (Thames Valley Test Company).

Byom, L. J., & Mutlu, B. (2013). Theory of mind: mechanisms, methods, and new directions. *Frontiers in Human Neuroscience*, *7*. https://doi.org/10.3389/fnhum.2013.00413

Canty, A. L., Neumann, D. L., Fleming, J., & Shum, D. H. K. (2017). Evaluation of a newly developed measure of theory of mind: The virtual assessment of mentalising ability. *Neuropsychological Rehabilitation*, *27*(5), 834–870. https://doi.org/10.1080/09602011.2015.1052820

Champagne-Lavau, M., & Moreau, N. (2013). Characteristics of current tasks that contribute to mentalizing judgments: Does the engagement of the participants in the social interaction matter? Comment on Achim et al. (2013). *Psychological Assessment*, *25*(4), 1404–1406. https://doi.org/10.1037/a0033433

Dahlberg, C., Hawley, L., Morey, C., Newman, J., Cusick, C. P., & Harrison-Felix, C. (2006). Social communication skills in persons with post-acute traumatic brain injury: Three perspectives. *Brain Injury*, *20*(4), 425–435. https://doi.org/10.1080/02699050600664574

de Bruin, L., van Elk, M., & Newen, A. (2012). Reconceptualizing second-person interaction. *Frontiers in Human Neuroscience*, *6*. https://doi.org/10.3389/fnhum.2012.00151

De Jaegher, H., Di Paolo, E., & Gallagher, S. (2010). Can social interaction constitute social cognition? *Trends in Cognitive Sciences*, *14*(10), 441–447. https://doi.org/10.1016/j.tics.2010.06.009

Delis, D. C., Kaplan, E., & Kramer, J. H. (2001). *Delis-Kaplan Executive Function System®(D-KEFS®): Examiner's Manual: Flexibility of Thinking, Concept Formation, Problem Solving, Planning, Creativity, Impluse Control, Inhibition*. Pearson.

Drapeau, J., Gosselin, N., Peretz, I., & McKerral, M. (2017). Emotional recognition from dynamic facial, vocal and musical expressions following traumatic brain injury. *Brain Injury*, *31*(2), 221–229. https://doi.org/10.1080/02699052.2016.1208846

Ekman, P., & Friesen, W. V. (1976). Measuring Facial Movement. *Environmental Psychology and Nonverbal Behavior*, *1*(1), 56–75.

El Haj, M., Raffard, S., & Gély-Nargeot, M.-C. (2015). Destination memory and cognitive theory of mind in normal ageing. *Memory*, 1–9. https://doi.org/10.1080/09658211.2015.1021257

Engberg, A. W., & Teasdale, T. W. (2004). Psychosocial outcome following traumatic brain injury in adults: a long-term population-based follow-up. *Brain Injury*, *18*(6), 533–545. https://doi.org/10.1080/02699050310001645829

Fazaeli, S. M., Amin Yazdi, S. A., Sharifi, S., Sobhani-Rad, D., & Ehsaei, M. R. (2018). Theory of Mind in Adults with Traumatic Brain Injury. *Trauma Monthly*, *23*(4). https://doi.org/10.5812/traumamon.22022

Frith, C. D. (2012). The role of metacognition in human social interactions. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *367*(1599), 2213–2223. https://doi.org/10.1098/rstb.2012.0123

Godefroy, O., Groupe de réflexion sur l'évaluation des fonctions, & exécutives. (2008). *Fonctions exécutives et pathologies neurologiques et psychiatriques.* Solal.

Hawthorne, G., Gruen, R. L., & Kaye, A. H. (2009). Traumatic Brain Injury and Long-Term Quality of Life: Findings from an Australian Study. *Journal of Neurotrauma*, *26*(10), 1623–1633. https://doi.org/10.1089/neu.2008.0735

Hynes, C. A., Stone, V. E., & Kelso, L. A. (2011). Social and emotional competence in traumatic brain injury: New and established assessment tools. *Social Neuroscience*, *6*(5–6), 599–614. https://doi.org/10.1080/17470919.2011.584447

Kelley, J. F. (1984). An iterative design methodology for user-friendly natural language office information applications. *ACM Transactions on Office Information Systems*, *2*, 26–41.

Kelly, M., McDonald, S., & Kellett, D. (2014). Development of a novel task for investigating decision making in a social context following traumatic brain injury. *Journal of Clinical and Experimental Neuropsychology*, *36*(9), 897–913. https://doi.org/10.1080/13803395.2014.955784

Kinsella, G., Packer, S., & Olver, J. (1991). Maternal reporting of behaviour following very severe blunt head injury. *Journal of Neurology, Neurosurgery, and Psychiatry*, *54*(5), 422–426.

Koskinen, S. (1998). Quality of life 10 years after a very severe traumatic brain injury (TBI): the perspective of the injured and the closest relative. *Brain Injury*, *12*(8), 631–648.

Marquardt, T. P., Rios-Brown, M., Richburg, T., Seibert, L. K., & Cannito, M. P. (2001). Comprehension and expression of affective sentences in traumatic brain injury. *Aphasiology*, *15*(10–11), 1091–1101. https://doi.org/10.1080/02687040143000429

Martin-Rodriguez, J. F., & León-Carrión, J. (2010). Theory of mind deficits in patients with acquired brain injury: A quantitative review. *Neuropsychologia*, *48*(5), 1181–1191. https://doi.org/10.1016/j.neuropsychologia.2010.02.009

May, M., Milders, M., Downey, B., Whyte, M., Higgins, V., Wojcik, Z., Amin, S., & O'Rourke, S. (2017). Social Behavior and Impairments in Social Cognition Following Traumatic Brain Injury. *Journal of the International Neuropsychological Society*, *23*(5), 400–411. https://doi.org/10.1017/S1355617717000182

McDonald, S. (2013). Impairments in Social Cognition Following Severe Traumatic Brain Injury. *Journal of the International Neuropsychological Society*, *19*(03), 231–246. https://doi.org/10.1017/S1355617712001506

McDonald, S., Bornhofen, C., Shum, D., Long, E., Saunders, C., & Neulinger, K. (2006). Reliability and validity of The Awareness of Social Inference Test (TASIT): A clinical test of social perception. *Disability and Rehabilitation*, *28*(24), 1529–1542. https://doi.org/10.1080/09638280600646185

McDonald, S., & Flanagan, S. (2004). Social Perception Deficits After Traumatic Brain Injury: Interaction Between Emotion Recognition, Mentalizing Ability, and Social Communication. *Neuropsychology*, *18*(3), 572–579. https://doi.org/10.1037/0894-4105.18.3.572

Mcdonald, S., & Saunders, J. C. (2005). Differential impairment in recognition of emotion across different media in people with severe traumatic brain injury. *Journal of the International Neuropsychological Society*, *11*(4), 392–399. https://doi.org/10.1017/S1355617705050447

Mutlu, B., Duff, M., & Turkstra, L. (2019). Social-cue perception and mentalizing ability following traumatic brain injury: A human-robot interaction study. *Brain Injury*, *33*(1), 23–31. https://doi.org/10.1080/02699052.2018.1531305

Piolino, P., Coste, C., Martinelli, P., Macé, A.-L., Quinette, P., Guillery-Girard, B., & Belleville, S. (2010). Reduced specificity of autobiographical memory and aging: Do the executive and feature binding functions of working memory have a role? *Neuropsychologia*, *48*(2), 429–440. https://doi.org/10.1016/j.neuropsychologia.2009.09.035

Premack, D., & Woodruff, G. (1978). Does the chimpanzee have a theory of mind? *Behavioral and Brain Sciences*, *4*(1), 515–526.

Radice-Neumann, D., Zupan, B., Babbage, D. R., & Willer, B. (2007). Overview of impaired facial affect recognition in persons with traumatic brain injury. *Brain Injury*, *21*(8), 807–816. https://doi.org/10.1080/02699050701504281

Sarfati, Y., Hardy-Baylé, M. C., Besche, C., & Widlöcher, D. (1997). Attribution of intentions to others in people with schizophrenia: a non-verbal exploration with comic strips. *Schizophrenia Research*, *25*(3), 199–209.

Schilbach, L., Timmermans, B., Reddy, V., Costall, A., Bente, G., Schlicht, T., & Vogeley, K. (2013). Toward a second-person neuroscience. *Behavioral and Brain Sciences*, *36*(04), 393–414. https://doi.org/10.1017/S0140525X12000660

Stone, V. E., & Hynes, C. A. (2011). Real-world Consequences of Social Deficits: Executive Functions, Social Competencies, and Theory of Mind in Patients with Ventral Frontal Damage and Traumatic Brain Injury. In *The Oxford Handbook of Social Neuroscience* (Oxford University Press, pp. 455–476). Jean Decety, John T. Cacioppo.

Teasdale, G., & Jennett, B. (1974). Assessment of coma and impaired consciousness. *The Lancet*, *304*(7872), 81–84. https://doi.org/10.1016/S0140-6736(74)91639-0

Temkin, N. R., Corrigan, J. D., Dikmen, S. S., & Machamer, J. (2009). Social Functioning After Traumatic Brain Injury: *Journal of Head Trauma Rehabilitation*, *24*(6), 460–467. https://doi.org/10.1097/HTR.0b013e3181c13413

Turkstra, L. S. (2008). Conversation-based assessment of social cognition in adults with traumatic brain injury. *Brain Injury*, *22*(5), 397–409. https://doi.org/10.1080/02699050802027059

Weschler, D. (2011). *Echelle d'Intelligence de Weschler pour Adultes - Quatrième Edition*. ECPA.

Williams, C., & Wood, R. Ll. (2010). Impairment in the recognition of emotion across different media following traumatic brain injury. *Journal of Clinical and Experimental Neuropsychology*, *32*(2), 113–122. https://doi.org/10.1080/13803390902806543

Zupan, B., Babbage, D., Neumann, D., & Willer, B. (2014). Recognition of facial and vocal affect following traumatic brain injury. *Brain Injury*, *28*(8), 1087–1095. https://doi.org/10.3109/02699052.2014.901560

Zupan, B., & Neumann, D. (2016). Exploring the Use of Isolated Expressions and Film Clips to Evaluate Emotion Recognition by People with Traumatic Brain Injury. *Journal of Visualized Experiments*, *111*. https://doi.org/10.3791/53774

Table 1 :

| | | TBI Participants | HC Participants | *t* value | *p*-value |
|---|---|---|---|---|---|
| **Demographical data** | | | | | |
| Age (years) | | 38 (10.5) | 37.6 (11.2) | 0.189 | >.05 |
| Education (years) | | 12.5 (2.5) | 13.1 (2.9) | -0.713 | >.05 |
| **Executive Functions** | | | | | |
| TMT | TMT A Time | 43.23*** (30.49) | 20.23 (5.86) | 4.125 | < .0001 |
| | TMT B Time | 96.23*** (56.96) | 47.45 (15.77) | 4.595 | < .0001 |
| | TMT B-A Time | 53*** (33.86) | 27.23 (14.22) | 3.907 | < .0001 |
| Hayling | Automatic score | 5.23*** (0.92) | 6.00 (0.26) | -4.509 | < .0001 |
| | Inhibition score | 9.68*** (2.36) | 12.26 (1.00) | -5.611 | < .0001 |
| Digit spans | Forward | 8.84*** (2.16) | 11.19 (1.33) | -5.169 | < .0001 |
| | Backward | 8.94*** (2.37) | 11.45 (2.03) | -4.494 | < .0001 |
| | Ordered | 8.65*** (2.64) | 12.90 (1.54) | -7.763 | < .0001 |
| Fluency | Semantic | 33.30*** (8.76) | 48.48 (9.69) | -6.477 | < .0001 |
| | Phonemic | 30.10*** (8.99) | 43.19 (11.32) | -5.044 | < .0001 |
| | Alternated | 11.94*** (3.04) | 16.84 (3.03) | -6.352 | < .0001 |
| **Memory** | | | | | |
| Autobiographical Fluency | AF1 | 5.61*** (1.87) | 9.26 (2.07) | | |
| | AF2 | 5.84*** (2.73) | 11.03 (3.58) | | |
| | AF3 | 3.13*** (2.11) | 6.71 (3.19) | | |
| | AF4 | 6*** (2.83) | 12.65 (3.09) | | |
| Semantic memory (Information) | | 12.27*** (3.93) | 16.13 (3.77) | -3.917 | < .0001 |
| Destination Memory | | 17.90*** (3.26) | 21.84 (1.85) | -5.850 | < .0001 |

Table 1: Neuropsychological results of patients with TBI and controls. ***difference between groups with $p$ <.001.

| | | Emotion_EVICog | ToM_EVICog | Ekman faces test | CIT |
|---|---|---|---|---|---|
| TMT B-A time | | -0.447** | -0.370* | -0.213 | -0.427* |
| Hayling | Inhibition score | 0.456** | 0.194 | 0.364* | 0.348* |
| Digit span | Forward | 0.375* | 0.481** | 0.209 | 0.661** |
| | Backward | 0.215 | 0.478** | 0.275 | 0.454** |
| | Ordered | 0.413* | 0.628** | 0.358* | 0.517** |
| Fluency | Semantic | 0.462** | 0.529** | 0.468** | 0.650** |
| | Phonemic | 0.433* | 0.289 | 0.225 | 0.228 |
| | Alternated | 0.546** | 0.306 | 0.404* | 0.334* |
| Autobiographical fluency (total score) | | 0.694** | 0.443** | 0.423* | 0.689** |
| Semantic memory | | 0.503** | 0.422* | 0.510** | 0.474** |
| Destination memory | | 0.394* | 0.607** | 0.418* | 0.498** |

Table 2: Correlations between social cognition scores and executive functions and memory scores.

*Significant at the $p < .05$ level; **Significant at the $p < .01$ level

**Figure 1**

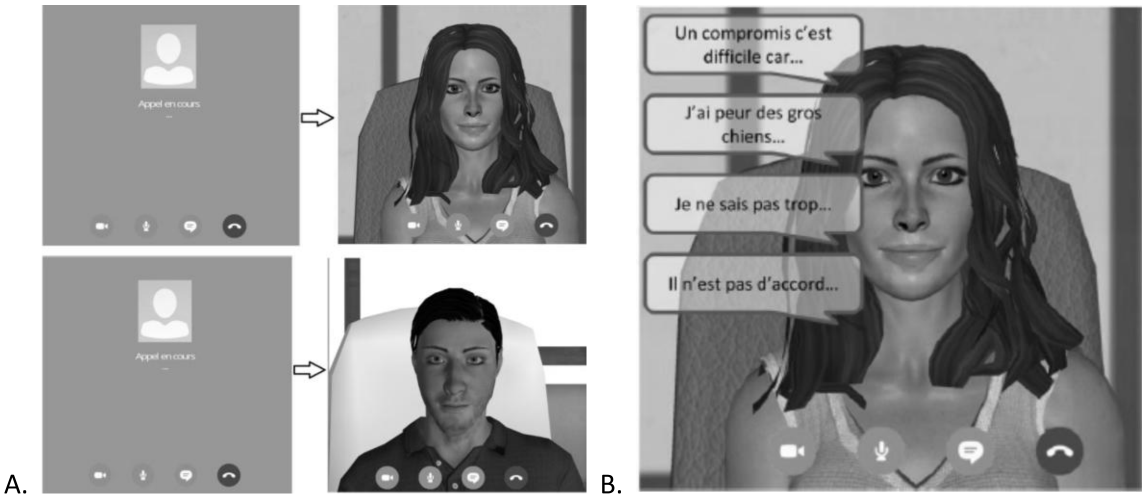
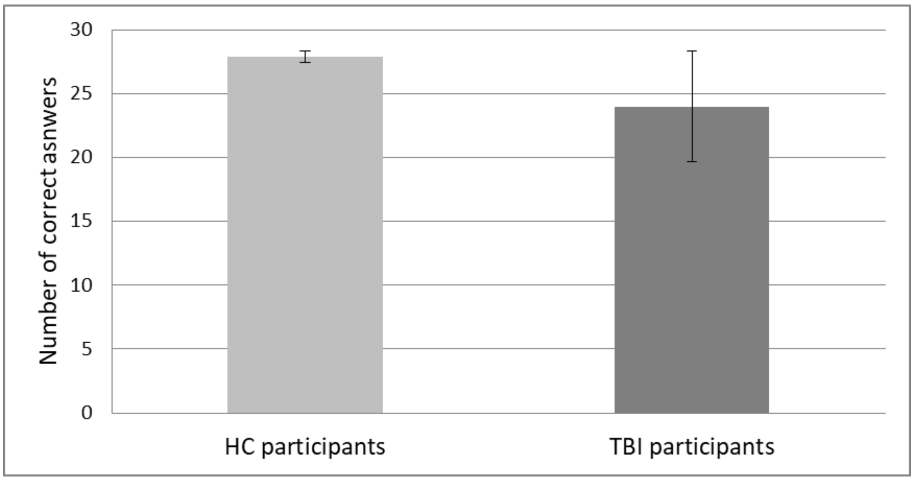



Figure 1. A: Participant's interface showing a conversation with Marie or Pierre. B : Experimenter's interface showing the different sentences the experimenter could select to answer the participant and continue the conversation.

**Figure 2**

A.




B.

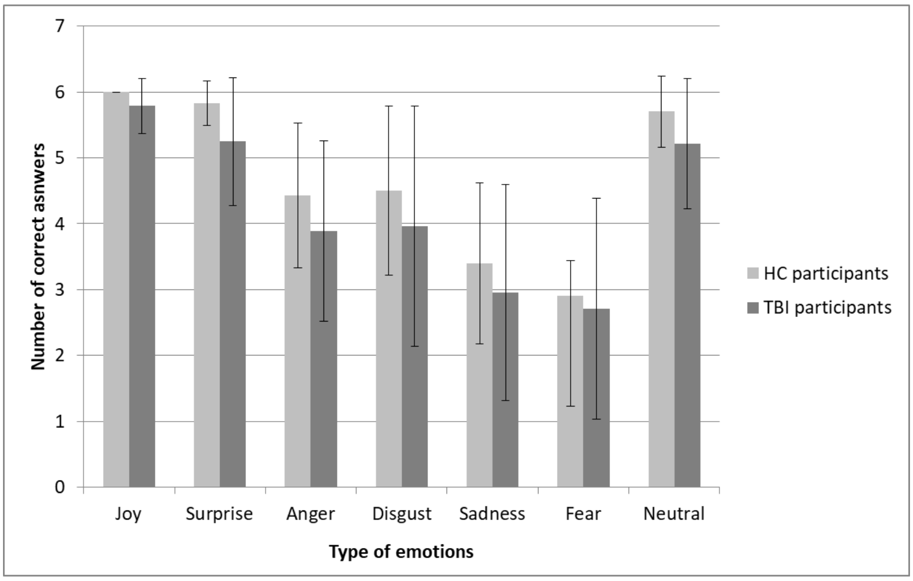



Figure 2. A: Total number of correct responses for each group on CIT. B: Number of correct responses for each group and each emotion on Ekman Faces test

**Figure 3**

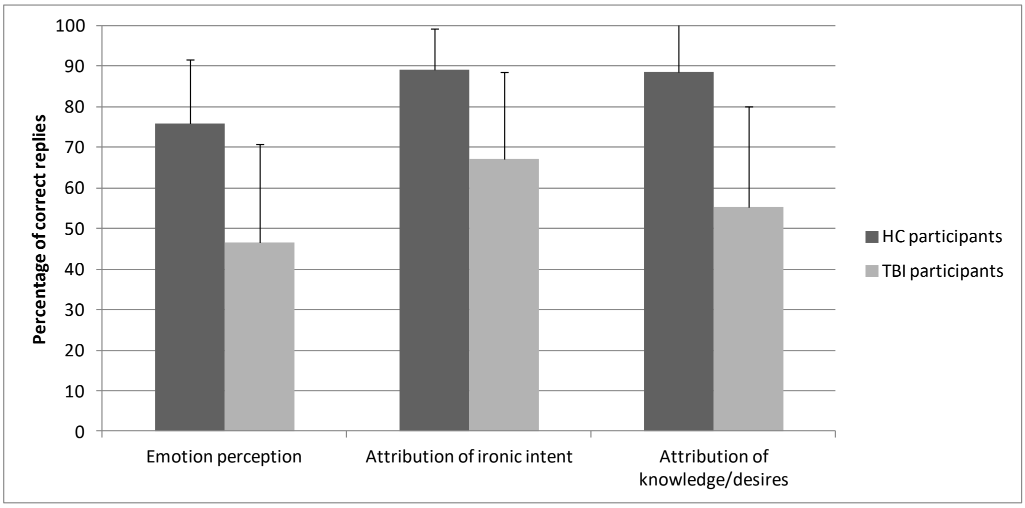



Figure 3. Percentage of correct replies produced in conversations for each domain of EVICog in both groups

**Figure 4**

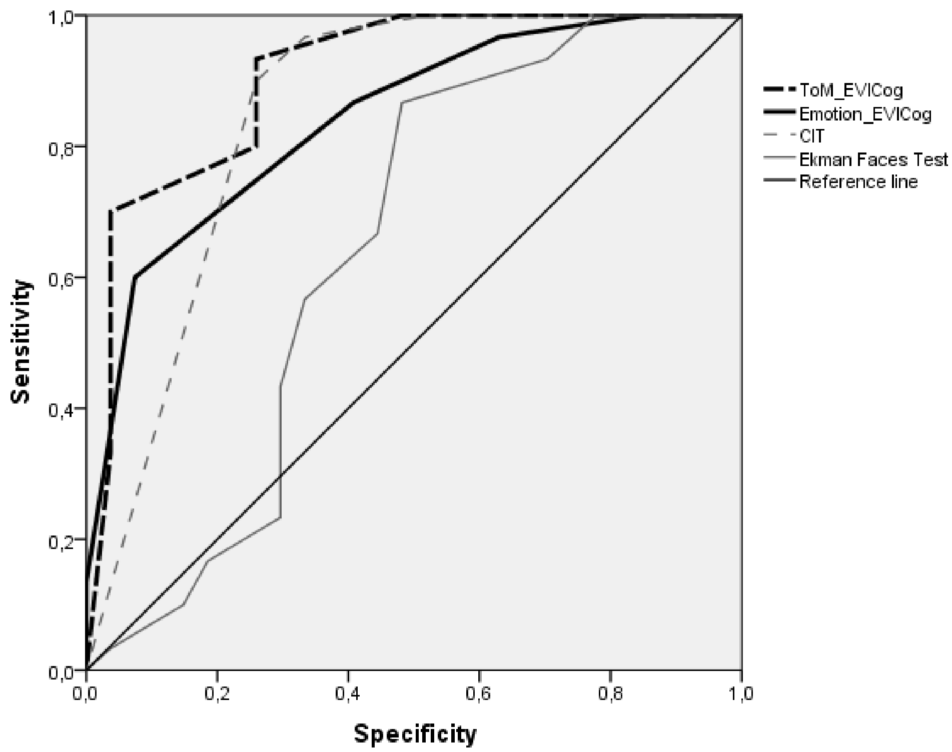



Figure 4. ROC curves for social cognition measures (ToM domain and emotion perception of EVICog, CIT, Ekman Faces test). The "reference line" reflects a test without discriminating power.