A virtual coach with more or less empathy: impact on older adults' engagement to exercise

Rachel Chauvin, Céline Clavel, Brian Ravenet, Nicolas Sabouret

A virtual coach with more or less empathy: impact on older adults' engagement to exercise

Rachel Chauvin
Université Paris-Saclay, LISN, CNRS
Orsay, France

Céline Clavel
celine.clavel@lisn.upsaclay.fr
Université Paris-Saclay, LISN, CNRS
Orsay, France

Nicolas Sabouret
nicolas.sabouret@lisn.upsaclay.fr
Université Paris-Saclay, LISN, CNRS
Orsay, France

Brian Ravenet
brian.ravenet@lisn.upsaclay.fr
Université Paris-Saclay, LISN, CNRS
Orsay, France

ABSTRACT

This paper presents a study that assesses the effectiveness of Motivational Interviewing (MI) interventions on the engagement of adults over 50 years to do more physical activity. While MI's four step process may easily be adapted to online devices, modeling the MI approach based on empathy may be more challenging (as opposed to directive interviews in which the user is encouraged to adopt a favorable attitude from the outset). Three devices are compared: a non-directive MI-based website, a non-directive MI-based virtual agent and a directed MI-based one. Our results show that the non-directive virtual agent tends to be perceived as more empathetic and trustworthy than the directed one,and that it significantly raises the participants' self-efficacy to overcome barriers, and positively impacts intrinsic motivation. We discuss mentions several implications for the adaptation of MI to the context of ECA.

CCS CONCEPTS

• **Human-centered computing → HCI design and evaluation methods**; **Empirical studies in HCI**; *Web-based interaction.*

KEYWORDS

conversational agent, motivational interviewing, physical activity, motivation, planning

ACM Reference Format:
Rachel Chauvin, Céline Clavel, Nicolas Sabouret, and Brian Ravenet. 2023. A virtual coach with more or less empathy: impact on older adults' engagement to exercise. In *ACM International Conference on Intelligent Virtual Agents (IVA '23), September 19–22, 2023, Würzburg, Germany.* ACM, New York, NY, USA, 9 pages. https://doi.org/10.1145/3570945.xxxxxxx

1 INTRODUCTION

1.1 Physical activity for health and ECAs

A large body of research has confirmed the detrimental effects of a sedentary lifestyle on health and psychological well-being on older adults, such as increased risks of cardiovascular and chronic diseases, mortality [22] [24]. Yet, the World Health Organization's global status report on physical activity (PA) 2022 alarmingly states that 27.5% of adults still do not meet the recommendation to do at least 150 minutes of moderate PA per week [24]. And adults over 50 years seem to have the lowest PA levels [22]. Many institutional prevention campaigns use text-based, multimedia websites. However, this type of informative content is not tailored to help individuals identify their personal needs and motivate themselves to act. One-to-one behavior change interventions with ECAs offer promising alternatives [16]. In their meta-analysis, Luo et al. (2021) noticed considerable variability in agent design and evaluation. They recommend to base future studies on theoretically sound approaches and on longitudinal assessments of the therapeutic effects [16].

1.2 Health behavior change approaches

Healthcare professionals tend to favor motivation-based interventions as they intend to empower the individual by raising their ability to recognize their internal motivators, modify their core beliefs, make reflected decisions and monitor their progress autonomously. They rely on motivation theories to convert the willingness to change into intention and action [8]. Motivation-based interventions differ from persuasive communication styles attempting to actively orientate attitudes and behaviors [7]. For instance, *positively framed counseling* techniques intend to direct the person's attention on positive thoughts that stimulate sustained positive emotions [28]. These strategies can be explained by the priming media effect which states that a repeated positive message should reinforce the association between the favorable attitude and the intention to change [6]. Such methods can certainly strengthen a person's decision to change in the absence of ambivalent beliefs, but they become ineffective when the recipient hold strong negative thoughts and is not ready to take action [19].

This research uses Motivational Interviewing (MI), a brief person-centered and empathetic consultation style focused on the enhancement of intrinsic motivation and self-efficacy to commit to change [18]. MI appears to be the most common intervention method used with virtual agents to promote PA [16].

1.3 Study overview

Despite the popularity of MI in human-virtual interactions, there is still a great variability in the aspects of MI exploited in ECA design

IVA '23, September 19–22, 2023, Würzburg, Germany

ACM ISBN 978-1-4503-9994-4/23/09.
https://doi.org/10.1145/3570945.xxxxxxx

[16]. Our objective was to conceive an ECA intervention suitable for French adults over 50 years who need to increase their PA level for health purposes. The scenario had to be aligned with MI's benevolent philosophy and grounded in evidence-based research and motivation theories. A mixed design study was implemented to evaluate how the virtual intervention could enhance older adults' motivation towards exercising and their engagement to do more PA the month following the session. To verify the extent to which the interaction with the ECA constitutes a valuable source of social presence capable of eliciting emotions and motivation through individualized guidance, we conceived a text-based website with the same informational content. Furthermore, empathy and respecting the client's state of readiness to change are crucial factors of MI's efficiency in human interactions [18]. We compared the effects of the empathic ECA with a more persuasive version forcing the user to think positively, that we named directed MI.

The following section reviews the theoretical approaches which led to the conception of the intervention and the study protocol.

2 RELATED WORK

2.1 Motivation and engagement to exercise

Due to the aging process and sedentary lifestyle of the population, various approaches are proposed to encourage the elderly to engage in more physical activity. These methods generally take into account the changing physical abilities of older individuals (such as loss of muscle mass, decreased flexibility, or balance issues), individual preferences (gentler exercises like yoga or swimming, or more dynamic activities like dance), social needs, and raising awareness of the benefits of physical activity. As a result, different interventions are designed, offering tailored exercise programs, realistic goals, and progress monitoring while promoting social support. For example, this is manifested in systems that provide activity tracking through devices such as pedometers and fitness bracelets. However, technological barriers, lack of personalization, and short-term motivation reinforcement can pose obstacles to the adoption and regular use of these devices. Furthermore, increased PA levels in older adults are associated with higher levels of motivation and self-efficacy [22].

Internal motivation can be explained through the concept of autonomous forms of motivation developed in self-determination theory (SDT) [32]. SDT stipulates that an individual will be much more inclined to undertake and pursue an activity if he feels driven by autonomous forms of motivation [32]. Intrinsic motivation, the propensity to act out of sheer pleasure or for the satisfaction of obtaining the results sought from exercising, represents the highest level of autonomy. At a lower level of freewill, extrinsic autonomous forms of motivation reflect a personal implication due to the identification and internalization of social values and norms into their self (integrated regulation)[3]. On the contrary, people who experience high levels of controlled forms or motivation feel constrained to act upon social pressures to meet social expectations (e.g., guilt), avoid punishment or seek rewards [32]. They will be less likely to maintain PA over time. Finally, amotivation represents the total absence of motivation.

Successful ECA interventions rely on strategies that raise the level of self-efficacy, that is the confidence in one's ability to pursue a specific course of action [16] [31]. According to the Heath Action Process Approach (HAPA), two forms of self-efficacy intervene in the initial engagement process [31]: (a) action self-efficacy, the confidence in the ability to start perform the task, which influences intention to act in conjunction with outcome expectations and risk perception; (b) barrier self-efficacy, the perceived ability to overcome obstacles, which supports action planning and coping strategies. Findings have shown that applying self-efficacy and goal-planning strategies jointly tend to significantly stimulate PA engagement [13] [27].

2.2 Motivational Interviewing premises

Motivational Interviewing (MI) is a brief intervention aimed at behavior change in one or two sessions [18]. It comprises two intertwined components: the relational component, emphasizing partnership, acceptance, evocation, and compassion, and the use of accurate empathy aligned with Carl Rogers's vision [18]. The practitioner refrains from giving advice without the client's consent and instead facilitates the client in exploring their values and resolving ambivalence. MI relies on monitoring change-talk and leveraging it to enhance internal motivation and reinforce self-efficacy [18]. By establishing a collaborative and empathetic relationship, MI elicits intrinsic motivations and builds clients' confidence in making positive changes [18].

2.3 MI adaptation to ECAs

Recent advancements in virtual counseling systems have shown promise in behavior change interventions. These systems, such as OCVIC [15] and the automated intervention system developed by Olafsson et al. [23], use virtual agents to deliver motivational interventions and promote positive health behaviors. OCVIC's empathetic virtual character has demonstrated significant increases in users' intention to use the intervention compared to text-based systems alone. Olafson et al.'s system, based on the Motivational Interviewing approach, employs virtual agents to enhance motivation and confidence in physical activity and fruit/vegetable consumption. The study presented in [23] is similar to the research we present in this paper, except that we focused on the needs of real users (sedentary aged people) and we studied a the medium-term impact of the intervention. In addition, we compared an agent-based intervention with a website version of the MI, which allows us to draw some conclusion on the effect of embodiment in the intervention and participants' engagements.

2.4 Hypotheses

The following hypotheses were drawn:

H1: The MI intervention based on empathy should increase users' autonomous and intrinsic motivation, self-efficacy and engagement to practice PA significantly more than a directed intervention forcing them to think positively (directed MI condition).

H2: The empathic agent should be perceived as more empathic and trustworthy than the agent who directs the conversation.

H3: An ECA intervention should have a higher perceived usability and satisfaction than a text-based website (H3a), leading to a higher increase in autonomous and intrinsic motivation and PA level after the intervention (H3b).

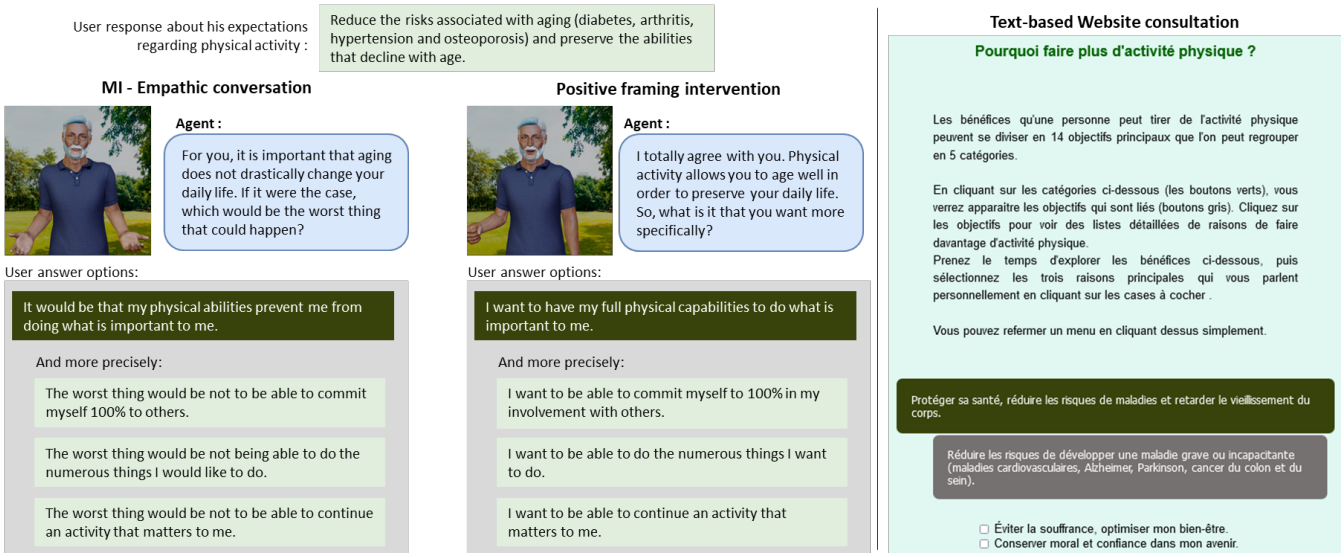

Figure 1: Extract of the MI scenario adapted the three experimental conditions.

You can access the interventions on `https://entretien-motivationnel.schm.fr` *(for the Empathic MI intervention),* `https://entretien-motivationnel2.schm.fr` *(for the directed MI intervention) and* `https://copains.lisn.upsaclay.fr/index.html` *(for the text-based website)*

3 INTERVENTION DESIGN

We first focused on the conception of the empathic conversation mimicking MI components as closely as possible subsection 2.2. The intervention was then adjusted to design a directed version of the agent and a text-based website (see Figure 1)

3.1 ECA interface

The computer-based interaction system followed a hierarchical structure with user-constrained input (option buttons and text boxes). To overcome the potential limitations of constrained speech, the dialogue contained a large number of pretested answer choices reflecting a variety of states of mind. The dialogue tree consisted of 81 statements, allowing for approximately 20 turns of conversation. At each turn, users were given a choice between 2 to 5 possible responses.

The virtual coach appearance was that of a 60 year-old, physically attractive and athletic man in sport-ware, called Jean (cf. Figure 1). Research has demonstrated on several occasions that an attractive human-like agent appeared more trustworthy and social than those who were not, in particular by older adults [30].

The synthesized voice and body gestures were carefully chosen to match the agent's age and profile. User acceptability depends on the fluency of the conversation and the congruence between verbal and nonverbal behavior [5]. An agent designed for counseling should have a low tone and relatively slow rhythm of voice, maintain eye contact (eye-sight facing the screen), lean the trunk slightly forward and reflect an open attitude with his face and arms [35] [25]. In this study, the agent's congruent behavior is all the more important that seniors seem to perceive virtual agents expressing emotions through nonverbal expressions with increased sensitivity compared to young adults [9].

3.2 Empathic MI intervention

3.2.1 Scenario structure. The intervention dialog followed the four recommended MI processes [18]. The first step, engagement, consisted in setting up a collaborative exchange. The agent invited the user to clarify their thoughts about PA and present the intervention objective. A pre-screening of the person's state of mind was done by asking them to rate to what extent they feel that PA is important on a scale from 1 to 10.

Next, during the evocation process users investigated their negative and positive beliefs. To elicit change-talk and thus boost their intrinsic motivation, the user was gradually driven to become aware of his needs and values. The dialogue was built around five pretested categories of objectives to engage in PA: health risks reduction, improvement of the body's appearance, psychological well-being, socialization, enhancement of the quality of life (pain reduction, sleep improvement). In the end, users could select three different reasons for doing PA among a total of 81 statements.

The reinforcement of self-efficacy process followed evocation. Users were asked to reflect upon the main difficulty likely to prevent them from taking action. The dialogue included five types of barriers identified by Justine (2013) [10]: not having enough time, social concerns like loneliness, limitations due to physical health issues, lack of motivation related to feelings of tiredness or laziness to practice, and practical concerns related to the difficulty to find a suitable activity near home. The agent then proposed to watch a video with testimonies of similar others who overcame successfully that specific difficulty. Such vicarious experiences are extremely beneficial to develop coping strategies that enhance self-efficacy, as the person may gain valuable insights that he would not have previously considered [1] [11].

Finally, users were carefully guided to choose an activity based on their personal preferences at the planning step. They could

select the goal of the activity (endurance, muscle-strengthening, balance, flexibility) and whether they prefer indoor or outdoor activities, practicing alone or in group. To strengthen the intention to engage in PA, an email follow-up in the name of the coach Jean was sent within 24 hours after the intervention. It contained a summary of the conversation and an action plan to fill-in at their own convenience during the week.

3.2.2 Amplification of empathic cues. The agent was conceived to mimic a human MI practitioner's accurate empathic attitude as closely as possible. The dialogue thus included regular reformulations and summaries to mirror users responses and provide gradual change-talk cues. To respect the person's readiness to change, the agent asked permission to continue the conversation at the end of each process [15]. The congruence between the agent's emotional expressions and user responses was controlled, taking into account the higher responsiveness of older adults to socio-emotional cues [9]. For example, the agent showed a sorry face and elbows raised if the user's response reflected fear [25]. Hosseinpanah et al. (2018) found that older adults considered that an agent was more empathic if he smiles and nod his head, or displays a sad face [9].

3.3 Directed MI intervention

The scenario was adjusted to infer a priming effect that would constrain the user to focus on the positive aspects of the intervention, so as stimulate extrinsic types of motivation. Based on Olafsson & al.'s (2019) coerced change-talk strategy [23], negatively framed user response statements were reformulated into positive ones.

Similarly, the agent's discourse was more optimistic, not giving time to the user to express self-doubts. Therefore, the evocation process became an opportunity for the user to focus only on the benefits of exercising, without thinking about what might happen if they didn't change. Figure 1 presents an extract of the conversation after a user answers that he wishes to reduce health risks associated with aging from the standpoint of an empathic MI and directed MI. The dialogue structure also encouraged the person's commitment to action without considering the state of readiness. All intermediary exit buttons were removed, forcing the user to finish the conversation and choose a practice. Affirmative statements (such as "here we go") replaced the agent's requests for the user's consent to go on. The agent's body language was adjusted to comply with a more persuasive dominant style [7]: a larger smile, more ample arm movements.

3.4 Text-based website

The multimodal website was designed to allow the user to read the content at his own pace. The interface was composed of 5 pages that followed the four-step scenario. Three short videos (about 10 seconds) of the agent welcomed users and congratulated them at the end of the session. Users had to use clickable links and buttons to progress through the website. They could freely navigate backwards only when they had already consulted a page. All vicariant videos from the reinforcement process were made available. Finally, the readiness stage was symbolized with an exit link at the end of each page.

4 STUDY IMPLEMENTATION

4.1 Participants and procedure

The participants were native French adults aged 50 years and older, without any major health issues preventing them from engaging in PA. Volunteers who expressed a feeling of not being active enough for their health were accepted. The 48 participants were recruited in France through social networks, cultural associations, associations of retirees and retirement homes. Among the 70.83% women and 29.17% men of mean age 64.11 (SD = 9.84), 64.58% were 60 or older (M = 69.90; SD = 7.57) and 60.42% were retired.

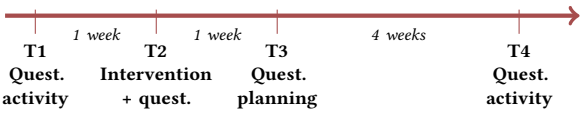

Figure 2: Protocol summary - From time T1 to T4

Participants were randomly assigned to one of the three web applications thanks to selection software (16 in each condition). The data collection process was divided in four steps over a month, summarized in Figure 2. First (T1), subscribers to the study completed an online questionnaire regarding their current PA level, motivation and self-efficacy. Second (T2), participants followed the online consultation which was held in the presence of a researcher either face-to-face (n = 33) or remotely through videoconferencing (n = 15). Upon the intervention completion, participants filled in a user acceptability survey. A follow-up email on behalf of the coach was sent two days after the consultation. It contained a summary of the coaching session along with a PA planning form as a guide to set up a new practice. Third (T3), an invitation to respond to email a new questionnaire was sent seven days after the consultation. It included motivation and self-efficacy scales, as well as behavioral questions about the planning stage of a PA. Finally (T4), participants were invited to fill-in the last questionnaire concerning the actions taken and their PA level four weeks after the coaching session. Forty-two participants completed all the steps.

4.2 Measures

4.2.1 Repeated measures. The following scales were administered before the intervention (T1) and after (T3 or T4) (cf. Figure 2).

Participants' PA level was assessed with the Global Physical Activity Questionnaire (GPAQ) endorsed by the World Health Organization [33]. It enables to calculate the average energy expenditure in terms of metabolic equivalent units (MET) minutes per week and classify the intensity of PA in three categories: low (e.g., slow walking), moderate (e.g., cycling) or high (e.g. playing football). One GPAQ response was omitted due to inconsistencies in the PA time reported at T1 and a four-month delay to answer at T4.

The motivation scale towards health-oriented physical activity measured the six types of self-determined regulation with 3-item subscales rated on a scale from 1 (*does not correspond at all*) to 7 (*corresponds very strongly*) [3]. For the analyses, the mean scores of intrinsic motivation and autonomous extrinsic motivation (composed of integrated regulation and identified regulation) were retained.

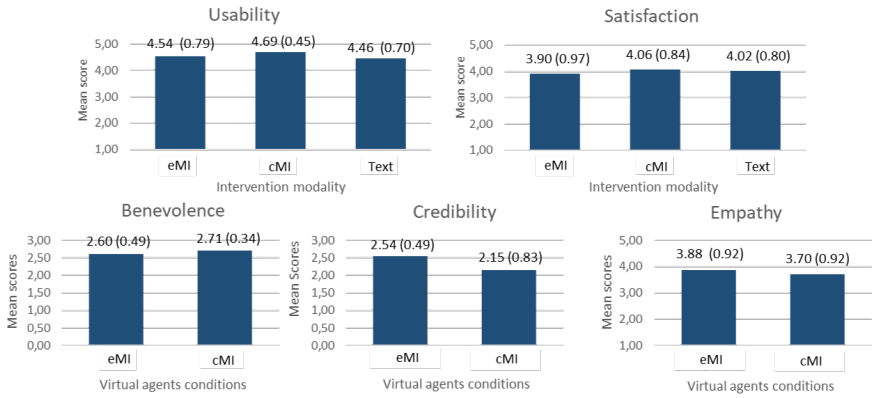

Figure 3: Acceptability and trust scores by modality.

Note: Sub-samples size n = 16, unless otherwise noted. eMI = empathic MI; cMI = directed MI; Text = text-based website; ns = non-significant result. Mean (SD) indicated above the bars. Independent t-tests and Mann-Witney U tests to compare conditions were all non-significant (p = n.s.).

Self-efficacy was rated with scores of certainty from 0 ('no certainty at all') to 10 ('total certainty'). Action self-efficacy measured with two items the perceived ability to do more PA in general and to meet the recommendations of the World Health Organization. Barrier self-efficacy (10 items) evaluated the perceived capability to overcome difficulties related to the choice of appropriate activity, time and physical constraints, lack of social support and motivation. The scale elaboration respected Bandura's guidelines [2] and was inspired by the self-efficacy for exercise literature [14] [29] [34].

4.2.2 Post-consultation attitudes (T2). The French version of the Acceptability e-scale (AES) [17] was used in all conditions to rate perceived usability (3 items) and satisfaction (3 items) on a 5-point Likert scale. The ECA Trust Questionnaire assessed users' perception of trust in the agent [26]. It contains a credibility and benevolence sub-scale of 3 items each rated from 0 ('not at all') to 3 ('totally agree'). Empathy was evaluated with the French version of the Consultation and Relational Empathy questionnaire (CARE) [21]. It is composed of 10 items rated from 1 ('poor') to 5 ('excellent'). As recommended, the CARE score was removed from two observations with more than two 'not applicable' responses [21].

Participants were also asked if they planned activities after the session (T3) and started a new PA one month later (T4).

4.2.3 Indicators from the consultation. In the empathic MI and text-based conditions, user readiness to change was assessed with a scale from 1 to 5 reflecting the completion level of the consultation: dropping out (1) after the agent's introduction, (2) after the evocation of needs, (3) before the planning phase without the intention to continue, (4) after the planning phase with the intention to plan an activity without assistance; (5) no dropping out. The time spent on the vicariant videos was also recorded.

5 RESULTS

Statistical analyses were conducted with Jasp V0.16.4. The psychometric attitudinal sub-scales' reliability estimated with Cronbach's alpha were all above the recommended .70 threshold, except for the self-efficacy action scale composed of only two items (T1: α = .63 ; T3 α = .61) [4]. Owing to the small sample (N < 50), The variables' potential deviation from normality was checked with the Shapiro-Wilk test [20]. Subgroup comparisons werre implemented with parametric and non-parametric t-tests (Mann-Withney U and Wilcoxon signed-rank) accordingly. Levene's test for equality of variances between subgroups showed no homogeneity issue between subgroups (p = n.s.).

5.1 Sample characteristics (T1)

No significant difference was observed between participants of the three conditions regarding mean age, gender, initial PA level, and number of retirees ($p > 0.10$) (cf. Table 3 in annex). Comparing the effects between conditions was therefore possible. Nevertheless, we acknowledged the large variation in PA levels owing to the random allocation of volunteers. The text-based website showed a high mean MET minutes of 2718.75 (SD = 3128.05) while having the highest number of participants with low PA level (n = 8) compared to each agent condition (n = 5).

The relatively high mean scores of intrinsic motivation (M = 4.58 ; SD = 1.46) and autonomous motivation (M = 4.86 ; SD = 1.29) indicated a relatively good perception of autonomy towards PA. Also, lower scores of barrier self-efficacy (M = 4.75 ; SD = 1.49) compared to action self-efficacy (M = 6.59 ; SD = 2.00) suggested some lack of confidence in overcoming difficulties to do PA.

5.2 Appreciation of the scenario

5.2.1 A widely appreciated scenario. When considering all conditions, participants appreciated the intervention, with high mean scores of satisfaction (M = 3.99; SD = 0.86) and usability (M = 4.56; SD = 0.65), the maximum score being 5 (cf. Figure 3). Only one participant in the empathic MI condition interrupted the conversation at the planning stage, preferring to postpone it and fill the planning form sent by email later. The video viewing patterns showed that 7 participants did not watch any vicariant experience (3 in the text-based website, 2 in each ECA condition). However, 5 users of the text-based website watched several videos (which was not possible in the agent conditions). Participants who watched videos

completely were significantly more satisfied (M = 4.16; SD = 0.71; Mdn = 4.33; n = 38) than those who didn't or only partially (M = 3.37; SD = 1.09; Mdn = 3.50; n = 10) (U = 109.00, p < .05, r_{pb} = -.43), suggesting that less satisfied users tended to skip the videos.

Findings tend to invalidate the hypothesis H3a. Comparisons between modalities showed no significant difference in usability and satisfaction between the empathic and directed MI conversation on one hand, and the text-based website on the other hand (p = ns) (cf. Figure 3). Perceived utility means scores of the empathic (M = 4.54; SD = 0.79) and directed agent (M = 4.69; SD = 0.45) were only slightly higher than the text-based website (M = 4.46; SD = 0.97) than Satisfaction was even slightly lower in the empathic condition (M = 3.90; SD = 0.97) than in the text-based website (M = 4.02; SD = 0.80) and directed MI (M = 4.06; SD = 0.84).

5.2.2 Reflection of trust and empathy in both agents. Benevolence and credibility mean scores above 2 over 3 revealed a good level of trust in both ECA interventions (cf. Figure 3). Similarly, mean empathy scores were over 3.5 over 5. Hypothesis H2 stating that user would find the empathic MI more trustworthy and empathic cannot be confirmed as no significant difference was found between the empathic MI and directed MI (p = n.s.). Directed MI (M = 2.71; SD = 0.34) was found almost more benevolent than empathic MI (M = 2.60; SD = 0.34). However, empathic MI (M = 2.54; SD = 0.49) tended to be perceived as more credible than directed MI (M = 2.15; SD = 0.83). Similarly, mean empathy scores were slightly higher with empathic MI (M = 3.88; SD = 0.92) and directed MI (M = 3.70; SD = 0.92).

5.3 Impact on attitudes and behavior

Prior to the analysis, the evolution of attitudes and PA level over time was calculated by subtracting the variables' scores after the intervention (T3 or T4) from the scores before the intervention (T1) (cf. Table 1, Table 2, Figure 4). A statistically positive difference indicates an improvement in attitude or PA level.

5.3.1 Mitigated global positive outcomes. A non-significant increase in intrinsic motivation, autonomous motivation, and self-efficacy was observed in the whole sample (p = n.s.). Out of the 44 respondents at T3, 24 (54.55%) mentioned having planned a new activity. Among them only 8 participants reported having implemented it a month later (T4). Nevertheless, we found a small but significant increase in total PA measured in MET-minutes between T1 and T4 (MD = 917.56; SD = 2453.78; W = 433.50, p < .05, r = 0.46).

5.3.2 Differentiated effect of modalities. Within subgroup paired-tests with each type of intervention showed no significant evolution in attitude and behavior for directed MI and text-based website conditions (cf. Figure 4, Table 1, Table 2). Negative differences indicate a tendency to reduced intrinsic and autonomous motivation. Regarding the empathic MI, barrier self-efficacy scores increased significantly (MD = 0.65; SD = 1.16; $t(14)$ = 2.17, p < .05, d= 0.56). And the increase was almost significant for intrinsic motivation (MD = 0.38; SD = 0.70; $t(14)$ = 2.09, p = 0.055, d = 0.54) and total PA level (MD = 1033.88; SD = 1849.03; W = 39.00, p = 0.057, r = 0.73).

Figure 4 highlights that the empathic MI tended to show greater increases in barrier self-efficacy, intrinsic and autonomous motivation and PA level than the other two conditions. Nevertheless,

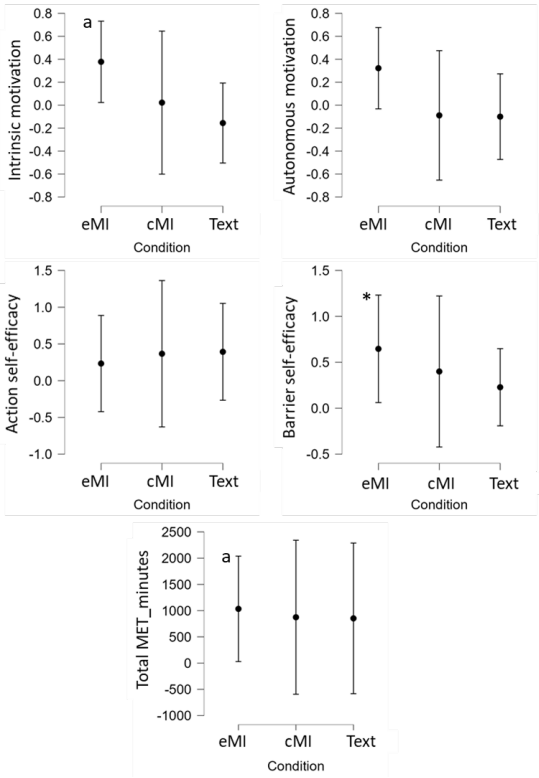

Figure 4: Evolution of motivation, self-efficacy and physical activity intensity among the three conditions.

Note: Evolution in terms of mean difference between scores before (T1) and after the session (T3, T4). Interval plots with 95% confidence intervals. eMI = empathic MI; cMI = directed MI; Text = text-based website. *: indicates a significant difference in that condition; a: indicates a p value between .05 and .06.

the increase in action self-efficacy scores was slighlty lower for the empathic MI (MD = 0.23; SD = 1.29) than the directed MI (MD = 0.37; SD = 1.97) and the text-based website (MD = 0.39; SD = 1.26).

However, no significant difference in attitudes and PA evolution was found between the empathic MI and directed MI (p = n.s.). Hypothesis H1 cannot be validated. Similarly, hypothesis H3b about the greater effectiveness of the empathic MI over the text-based website cannot be confirmed. The only significant difference was regarding the higher evolution of intrinsic motivation for the empathic MI intervention (MD = 0.38; SD = 0.70) compared to the website consultation (MD = -0.16; SD = 1.16) ($t(28)$ = 2.10, p < .05, d= 0.77).

These findings should be considered with caution, owing to the high individual variations reflected in the confidence intervals.

6 DISCUSSION

6.1 Main findings

This study explored the extent to which a single ECA intervention integrating most of MI features could foster older adults' motivation and engagement to undertake more physical activity (PA).

Table 1: Evolution of attitudes and physical activity for the sample and the empathic condition.

	Whole sample			Empathic MI		
Variable	n	Difference MD(SD)	Paired test - effect size (d or r)	n	Difference MD(SD)	Paired test - Effect size (d or r)
Intrinsic motivation	45	0.081 (0.92)	t(44) = 0.60, p = n.s., d = 0.089	15	0.38 (0.70)	t(14) = 2.09, p = 0.055, d = 0.54
Autonomous motivation	45	0.044 (0.87)	t(44) = 0.34, p = n.s., d = 0.051	15	0.32 (0.70)	t(14) = 1.78, p = n.s., d= 0.46
Task self-efficacy	44	0.33 (1.513)	t(43) = 1.45, p = n.s.; d = 0.22	15	0.23 (1.29)	t(14) = 0.70, p = n.s., d= 0.18
Barrier self-efficacy	44	0.43 (1.232)	Mdn = .30; W = 644.50, p < .01, r = 0.50	15	0.65 (1.16)	t(14) = 2.17, p < .05, d= 0.56
Total MET-minutes	41	917.56 (2453.78)	Mdn = 360; W = 433.50, p < .05, r = 0.46	13	1033.85 (1849.03)	Mdn = 480; W = 39.00, p = .057, r = 0.73

Note: Mean difference MD = score after (T3 or T4) minus score before the intervention (T1). Wilcoxon rank test was used in place of the Student t-test when data deviated from normal distribution. Effect size measured with Cohen's d or r rank biserial correlation.

Table 2: Evolution of attitudes and physical activity for the directed MI and Website condition.

	Directed MI		**Website**	
Variable	n	Difference MD(SD)	n	Difference MD(SD)
Intrinsic motivation	15	0.022 (1.23)	15	-0.16 (0.69)
Autonomous motivation	15	-0.089 (1.11)	15	-0.10 (0.74)
Task self-efficacy	15	0.37 (1.97)	14	0.39 (1.26)
Barrier self-efficacy	15	0.40 (1.62)	14	0.23 (0.80)
Total MET-minutes	14	874.29 (2800.71)	14	852.86 (2739.78)

Note: Mean difference MD = score after (T3 or T4) minus score before the intervention (T1). T-tests were non-significant.

Findings suggest that the scenario based on the four MI processes was well appreciated, all experimental conditions confounded. However, there was no evidence of better perceived trust and empathy between the agents nor improved satisfaction among the three intervention modalities. The empathic MI obtained slightly higher scores of credibility and empathy than the positively framed interview, which is encouraging. The lower satisfaction scores with the empathic agent could be related to individual differences or the more complex dialog including readiness questions. Philip et al. (2020) argued that patients were less willing to engage in another interaction with an ECA for depression detection than with an ECA for addiction diagnosis because the dialog for depression contained more open-ended and reflective questions than the one for addictions [26].

Despite the lack of significant results when comparing modalities, the empathic agent led to larger increases in self-efficacy to overcome obstacles, intrinsic and autonomous motivation than a directed MI and the text-based website. The empathic MI had a significant positive effect on barrier self-efficacy but not on action self-efficacy and little on planning, which could explain the lack of commitment to do more PA in the long run [13]. Action self-efficacy is supposed to stimulate the intention to engage in PA, while barrier self-efficacy is more likely to facilitate planning and behavior regulation [31]. The videos full of advice to overcome difficulties may have triggered barrier-self-efficacy more than action self-efficacy.

Besides, a single intervention may not be enough to induce a healthy behavior. Most studies who reported a positive impact on PA initiation and adherence were generally based on repeated virtual interactions [12] [16]. This study is indeed part of a larger project with the intention to design a second MI dialogue focused on the planning phase and follow-up of PA. At this stage, the priority was to assess the value of the first intervention and make the necessary improvements. And the planning phase was only addressed by email.

6.2 Limitations and perspective

The small sample size in each condition prevented us from deepening statistical analyses. Because recruiting sedentary people was difficult (no compensation provided), any volunteer who thought they were not doing enough PA for health were accepted. In the end, only 18 participants reported a low initial PA level while 11 reported an intensive level due to work activities (non-profit organizations, household activities) and travel. Replication and extensions of the present research should consider a larger sample with a prescreening of PA levels. Another potential limitation was the use of self-report measures to assess PA. Some participants reported having difficulties to fill in the Global Physical Activity Questionnaire (GPAQ), even though the version was adapted to self-reports [33]. Real-time tracking devices to assess the evolution of PA should limit potential respondent biases. Finally, the manipulation of the scenario between the conditions may have been too subtle to generate significant differences. The positively framed agent was designed to focus on positive emotions. Our results suggest that some directed MI features could improve the empathic MI version without distorting MI's philosophy. For instance, some exit dialogues could be removed to lighten the scenario, as no user in the empathic MI dropped out before the planning stage. Olafsson et al. (2019) recommended such adjustments when they found that users were significantly more confident with the coerced change-talk condition than the non-coerced one [23].

7 CONCLUSION

The design of health interventions based on motivational interviewing should respect its core principles to preserve its integrity. Nevertheless, their adaptation to ECA systems may require flexibility as the importance of some criteria may differ from human interactions. Future work on physical activity for older adults could benefit from real-time measures and the study of the effect of individual characteristics such as the locus of control and subjective health beliefs that appear linked to older adults PA levels [22]. Another challenge may be to develop systems with personalized conversations adapted to multiple individual profiles.

ACKNOWLEDGMENTS

This research was funded by the French National Research Agency as part of the COPAINS project. The agent was designed by DAVI - The Humanizers, partner of the project.

REFERENCES

[1] Albert Bandura. 1997. *Self-efficacy: The exercise of control.* W H Freeman/Times Books/ Henry Holt & Co, New York, NY, US. Pages: ix, 604.

[2] Albert Bandura et al. 2006. Guide for constructing self-efficacy scales. *Self-efficacy beliefs of adolescents* 5, 1 (2006), 307–337.

[3] Julie Boiché, Mathieu Gourlan, David Trouilloud, and Philippe Sarrazin. 2019. Development and validation of the 'Echelle de Motivation envers l'Activité Physique en contexte de Santé': A motivation scale towards health-oriented physical activity in French. *Journal of Health Psychology* 24, 3 (March 2019), 386–396. Number: 3 Publisher: SAGE Publications Ltd.

[4] Jose M. Cortina. 1993. What is coefficient alpha? An examination of theory and applications. *Journal of Applied Psychology* 78 (1993), 98–104. Place: US Publisher: American Psychological Association.

[5] Ylva Ferstl, Sean Thomas, Cédric Guiard, Cathy Ennis, and Rachel McDonnell. 2021. Human or Robot? Investigating voice, appearance and gesture motion realism of conversational social agents. In *Proceedings of the 21st ACM International Conference on Intelligent Virtual Agents*. Association for Computing Machinery, New York, NY, USA, 76–83.

[6] Martin Fishbein and Marco C. Yzer. 2003. Using Theory to Design Effective Health Behavior Interventions. *Communication Theory* 13, 2 (2003), 164–183. Number: 2 _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1468-2885.2003.tb00287.x.

[7] Océane Granier, Roxane Bertrand, and Magalie Ochs. 2021. La persuasion dans les interactions humain-humain et humain-machine: une revue de la littérature. In *WACAI 2021*. Centre National de la Recherche Scientifique [CNRS], Saint Pierre d'Oléron, France, 1–9.

[8] Sarah J. Hardcastle, Jennie Hancox, Anne Hattar, Chloe Maxwell-Smith, Cecilie Thøgersen-Ntoumani, and Martin S. Hagger. 2015. Motivating the unmotivated: how can health behavior be changed in those unwilling to change? *Frontiers in Psychology* 6 (2015), 1–4.

[9] Adineh Hosseinpanah, Nicole C. Krämer, and Carolin Straßmann. 2018. Empathy for Everyone? The Effect of Age When Evaluating a Virtual Agent. In *Proceedings of the 6th International Conference on Human-Agent Interaction (HAI '18)*. Association for Computing Machinery, New York, NY, USA, 184–190. event-place: Southampton, United Kingdom.

[10] Maria Justine, Azliyana Azizan, Vaharli Hassan, Zoolfaiz Salleh, and Haidzir Manaf. 2013. Barriers to participation in physical activity and exercise among middle-aged and elderly individuals. *Singapore Medical Journal* 54, 10 (Oct. 2013), 581–586.

[11] Jeemin Kim, Mark Eys, and Jennifer Robertson-Wilson. 2021. 'If they do it, so can I': a test of a moderated serial mediation model of descriptive norms, self-efficacy, and perceived similarity for predicting physical activity. *Psychology & Health* 36, 6 (June 2021), 701–718. Publisher: Routledge.

[12] Rafal Kocielnik, Lillian Xiao, Daniel Avrahami, and Gary Hsieh. 2018. Reflection Companion: A Conversational System for Engaging Users in Reflection on Physical Activity. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 2, 2 (July 2018), 70:1–70:26. Number: 2.

[13] Milena Koring, Jana Richert, Sonia Lippke, Linda Parschau, Tabea Reuter, and Ralf Schwarzer. 2012. Synergistic Effects of Planning and Self-Efficacy on Physical Activity. *Health Educ Behav* 39, 2 (2012), 152–158. Place: Los Angeles, CA Publisher: Los Angeles, CA: SAGE Publications.

[14] Thilo Kroll, Matthew Kehn, Pei-Shu Ho, and Suzanne Groah. 2007. The SCI Exercise Self-Efficacy Scale (ESES): development and psychometric properties. *International Journal of Behavioral Nutrition and Physical Activity* 4, 1 (Aug. 2007), 34.

[15] Christine Lisetti, Reza Amini, Ugan Yasavur, and Naphtali Rishe. 2013. I Can Help You Change! An Empathic Virtual Agent Delivers Behavior Change Health Interventions. *ACM Transactions on Management Information Systems* 4, 4 (Dec. 2013), 19:1–19:28. Number: 4.

[16] Tiffany Christina Luo, Adrian Aguilera, Courtney Rees Lyles, and Caroline Astrid Figueroa. 2021. Promoting Physical Activity Through Conversational Agents: Mixed Methods Systematic Review. *Journal of Medical Internet Research* 23, 9 (Sept. 2021), e25486. Number: 9 Company: Journal of Medical Internet Research Distributor: Journal of Medical Internet Research Institution: Journal of Medical Internet Research Label: Journal of Medical Internet Research Publisher: JMIR Publications Inc., Toronto, Canada.

[17] Jean-Arthur Micoulaud-Franchi, Alain Sauteraud, Jérôme Olive, Patricia Sagaspe, Stéphanie Bioulac, and Pierre Philip. 2016. Validation of the French version of the Acceptability E-scale (AES) for mental E-health systems. *Psychiatry Research* 237 (March 2016), 196–200.

[18] William R. Miller and Stephen Rollnick. 2013. *Motivational interviewing: Helping people change, 3rd edition.* Guilford Press, New York, NY, US. Pages: xii, 482.

[19] William R. Miller and Gary S. Rose. 2015. Motivational Interviewing and Decisional Balance: Contrasting Responses to Client Ambivalence. *Behavioural and Cognitive Psychotherapy* 43, 2 (March 2015), 129–141. Number: 2 Publisher: Cambridge University Press.

[20] Prabhaker Mishra, Chandra M Pandey, Uttam Singh, Anshul Gupta, Chinmoy Sahu, and Amit Keshri. 2019. Descriptive Statistics and Normality Tests for Statistical Data. *Annals of Cardiac Anaesthesia* 22, 1 (2019), 67–72.

[21] Maxime Mons. 2015. *Empathie et odontologie.* exercice. Université Toulouse lll - Paul Sabatier.

[22] Nanna Notthoff, Peter Reisch, and Denis Gerstorf. 2017. Individual Characteristics and Physical Activity in Older Adults: A Systematic Review. *Gerontology* 63, 5 (2017), 443–459. Publisher: Karger Publishers.

[23] Stefan Olafsson, Teresa O'Leary, and Timothy Bickmore. 2019. Coerced Change-talk with Conversational Agents Promotes Confidence in Behavior Change. In *Proceedings of the 13th EAI International Conference on Pervasive Computing Technologies for Healthcare (PervasiveHealth'19)*. Association for Computing Machinery, New York, NY, USA, 31–40.

[24] World Health Organization. 2022. *Global status report on physical activity 2022.* Technical Report. World Health Organization, Geneva. 1–132 pages. 978-92-4-005915-3.

[25] Ana Paiva, Iolanda Leite, Hana Boukricha, and Ipke Wachsmuth. 2017. Empathy in virtual agents and robots: A survey. *ACM Transactions on Interactive Intelligent Systems (TiiS)* 7, 3 (2017), 1–40. Number: 3 Publisher: ACM New York, NY, USA.

[26] Pierre Philip, Lucile Dupuy, Marc Auriacombe, Fushia Serre, Etienne de Sevin, Alain Sauteraud, and Jean-Arthur Micoulaud-Franchi. 2020. Trust and acceptance of a virtual psychiatric interview between embodied conversational agents and outpatients. *npj Digital Medicine* 3, 1 (Jan. 2020), 2. Number: 1.

[27] Janani Pinidiyapathirage, Rohan Jayasuriya, N. Wah Cheung, and Ralf Schwarzer. 2018. Self-efficacy and planning strategies can improve physical activity levels in women with a recent history of gestational diabetes mellitus. *Psychology & Health* 33, 8 (Aug. 2018), 1062–1077. Publisher: Routledge _eprint: https://doi.org/10.1080/08870446.2018.1458983.

[28] Jordi Quoidbach, Moïra Mikolajczak, and James J. Gross. 2015. Positive interventions: An emotion regulation perspective. *Psychological bulletin* 141, 3 (2015), 655. Number: 3 Publisher: American Psychological Association.

[29] B. Resnick and L. S. Jenkins. 2000. Testing the reliability and validity of the Self-Efficacy for Exercise scale. *Nursing Research* 49, 3 (2000), 154–159.

[30] Minjin Rheu, Ji Youn Shin, Wei Peng, and Jina Huh-Yoo. 2021. Systematic Review: Trust-Building Factors and Implications for Conversational Agent Design. *International Journal of Human–Computer Interaction* 37, 1 (Jan. 2021), 81–96. Number: 1 Publisher: Taylor & Francis _eprint: https://doi.org/10.1080/10447318.2020.1807710.

[31] Ralf Schwarzer, Sonia Lippke, and Aleksandra Luszczynska. 2011. Mechanisms of health behavior change in persons with chronic illness or disability: The Health Action Process Approach (HAPA). *Rehabilitation Psychology* 56, 3 (2011), 161–170. Number: 3 Place: US Publisher: American Psychological Association.

[32] Robert J. Vallerand. 2007. Intrinsic and extrinsic motivation in sport and physical activity: A review and a look at the future. In *Handbook of sport psychology, 3rd ed.* John Wiley & Sons, Inc., Hoboken, NJ, US, 59–83.

[33] Miriam Wanner, Christina Hartmann, Giulia Pestoni, Brian Winfried Martin, Michael Siegrist, and Eva Martin-Diener. 2017. Validation of the Global Physical Activity Questionnaire for self-administration in a European context. *BMJ Open Sport & Exercise Medicine* 3, 1 (May 2017), e000206. Publisher: BMJ Specialist Journals Section: Original article.

[34] Tsu-Yin Wu, Lorraine B. Robbins, and Hsing-Fang Hsieh. 2011. Instrument development and validation of Perceived Physical Activity Self-Efficacy Scale for adolescents. *Research and Theory for Nursing Practice* 25, 1 (2011), 39–54.

[35] Bo Xiao, Daniel Bone, Maarten Van Segbroeck, Zac E. Imel, David C. Atkins, Panayiotis G. Georgiou, and Shrikanth S. Narayanan. 2014. Modeling therapist empathy through prosody in drug addiction counseling. In *Fifteenth annual conference of the international speech communication association*. International Speech Communication Association (ISCA), Singapore, 213–217. September 14-18, 2014.

A APPENDIX

Table 3: Sample distribution by modality of intervention

	Empathic MI	Directed MI	Website	Total (N = 48)	Chi Test
Age - M(SD)	64.93 (7.73)	65.75 (11.36)	61.53 (10.12)	64.11 (9.84)	$\chi^2(2, 48) = 5.10, p = 0.078$
60 and older	13	11	7	31	
Less than 60	3	5	9	17	
Gender					$\chi^2(2, 48) = 1.41, p = 0.49$
Male	10	11	13	34	
Female	6	5	3	14	
Status					$\chi^2(2, 48) = 2.23, p = 0.33$
Retired	11	10	7	28	
In activity	5	6	9	20	
Physical activity level (T1)					$\chi^2(4, 47) = 4.57, p = 0.33$
Low	5	5	8	18	
Moderate	8	7	3	18	
High	2	4	5	11	
Live session					$\chi^2(2, 48) = 0.58, p = 0.75$
Distance	4	5	6	15	
Face to face	12	11	10	33	