

# A Highly Elaborative Reminiscing Virtual Agent to Enhance Student Memory of Virtual World Events

Michelle Nicholas<sup>1</sup>, Deborah Richards<sup>2</sup>, Penny Van Bergen<sup>3</sup>

<sup>1</sup>Department of Psychology <sup>2</sup>Department of Computing <sup>3</sup>Department of Education

Macquarie University

North Ryde, NSW, 2109, Australia

<sup>2</sup>+61 2 9850 9567

{deborah.richards,penny.vanbergen}@mq.edu.au

## ABSTRACT

The use of “highly elaborative” reminiscing by mothers, involving open-ended questions and more detail about past events has been shown to produce children with greater accuracy and detail in their autobiographical memory recall of past events. This paradigm has not been extended to semantic learning, or to an adult population before. Thus, the present study used a virtual reminiscing agent to investigate the effect of elaborative and non-elaborative reminiscing on students’ semantic memory recall of a virtual event to investigate this reminiscing paradigm. A study involving ninety-eight undergraduate psychology students revealed that participants exposed to highly elaborative reminiscing free recalled significantly more correct information than low elaborative reminiscing (closed-ended questions and low detail), and no reminiscing. Second, open-ended questions were superior to closed-ended questions in free recall memory, as well as in number of memory errors. Lastly, the provision of high detail improved recall when combined with open-ended questions, but not when combined with closed-ended questions. These results indicate that an intelligent virtual agent with a highly elaborative reminiscing style may be a valuable learning companion by aiding (adult) students to reflect on and better remember episodic experiences provided in a virtual world.

## Categories and Subject Descriptors

I.2.11 [Artificial Intelligence]: Distributed Artificial Intelligence – intelligent agents, multiagent systems.

## General Terms

Design and Human Factors.

## Keywords

Reminiscing agent, semantic memory, episodic memory.

## 1. INTRODUCTION

Experiential Learning Theory (ELT) [17] highlights the importance of experience and reflection in any experiential learning process. Reflection is considered a critical part of learning at all stages of education: primary, secondary, tertiary, and professional [31]. Experiential virtual worlds (VWs) engage episodic and autobiographical memory processes because there is a sense of time and personal involvement for students involved in these worlds [22]. Strategies known to promote autobiographical memory in children might also be applied to experiences in a VW.

**Appears in:** *Proceedings of the 12th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2013)*, Ito, Jonker, Gini, and Shehory (eds.), May, 6–10, 2013, Saint Paul, Minnesota, USA. Copyright © 2013, International Foundation for Autonomous Agents and Multiagent Systems (www.ifaamas.org). All rights reserved.

This is important to investigate since there exists little research into how best to design pedagogies to promote reflection and learning in VWs [32]. Moreover, since episodic-spatial and semantic-spatial memories are activated similarly in the hippocampus, semantic learning outcomes (i.e., memory for facts and knowledge) can be implemented into a VW event [15].

While little educational research has examined the use of reflection in VWs, the developmental research does highlight the benefit of scaffolded reminiscing for children’s recall of personally experienced past events. The most effective form of scaffolding for children in this case is one that contains both open-ended questioning and high detail (i.e., high elaboration) [12]. However, this paradigm has not been extended to semantic learning, or to an adult population before. The value of open-ended questions has been investigated in educational psychology [18], but the benefits of “high detail” on semantic learning have not. The provision of extra detail during reflective reminiscing might conceivably lead to greater improvements in semantic knowledge acquisition. This extra level of elaboration would provide the student with more memory cues in which to better encode, consolidate and recall this information.

This paper concerns an evaluation of the potential benefits of using a Reminiscing Agent (RA) for student reflection and its impact on their learning. In particular, we seek to understand what reminiscing conversational style may be most beneficial for memory recall. Because an RA is virtual and not constrained by time, space, availability or budgets in the way that a human teacher, peer learner or companion may be, the use of an RA to improve episodic, autobiographical and semantic recall of content learnt in a VW could significantly change current practice.

Furthermore, for the purposes of this investigation, the use of an RA and a VW event overcame the problem of experimental control that would be difficult to achieve if seeking to assess the effectiveness of reminiscing talk styles in teacher to student classroom interaction. Using the RA, it was easier to control conversational output and thereby better differentiate what questioning styles were most effective for facilitating learning.

### 1.1 Agents, reflection and memory research.

Memory in virtual agents has typically been implemented to enable agents to remember information in order to carry out their intended purpose (games, narratives, assistive computing, pedagogical agents, etc.). Memory also improves the agent’s believability and builds better interpersonal relationships with the user. To achieve this, memory may be incorporated into emotion-appraisal models, as in the case of the virtual geography teacher, Eva [16]. Memory-related research can be found in many projects that are part of EU-funded projects LIREC and COMPANION

since long-term relationships with IVAs can't ignore memory [14] but need to address issues such as what companions need to remember and even what they should forget [36]. There is increasing interest in designing agents with (auto)biographical memory (e.g. [11]) which enable agents to dynamically reconstruct their individual life histories [1, 11]. Preliminary work on engagement and deceit [4] found that agents that tell autobiographical stories from the first person perspective were more engaging and more likely to encourage the user to continue to use the system.

To improve believability and usefulness, much of the work in IVAs draws upon the psychology-based memory literature seeking to replicate human memory processing in agents (e.g. [13, 20, 34]). Brom and Lukavský [5] offer a general framework for episodic memory in agents to support agents in a broad range of tasks like debriefing (similar to the task performed by our RA), giving information, remembering the course of interactions, searching for objects, and knowledge sharing and learning. Some research seeks to understand what humans expect in agent memory behaviours. For example, Burkert et al. [6] looked at the level of detail associated with episodic timing information and found that people prefer to use socially established time patterns rather than exact times or dates when questioning IVAs. Lim et al [20] found that users preferred "biological" over "permanent not erasable" and "permanent but erasable (ie reset)" memories. They also reported that users found selective memory more natural than absolute memory, but tended to prefer the companion with absolute memory. Vargas et al [38] are working on "roboethical" memory for robot companions, focusing on forgetting mechanisms to control the information stored and retrieved.

The MAY (My Memories Are Yours) agent [7] uses a conceptual model for shared memories comprising event-specific knowledge, general events and lifetime periods. An experiment using MAY found that when memories about the user were exhibited the scores for intimacy and companionship were higher. MAY is based on the theory that episodic memory is retained only when linked to autobiographical/semantic memory, consistent with our study. For this reason our RA shares the VW event/experience with the user, though our RA does not have its own memories.

Research directly in the reflection space includes the DARPA funded Reflective Agents with Distributed Adaptive Reasoning [28] project that embedded machine learning technology to allow a cognitive assistant to adapt and train itself dynamically without the need for a human expert. The REflective Agent Learning environment project (REAL) [2] involves students first creating an imaginary world, which they then reflect upon. These reflections are used by the reflective agent to dynamically generate a simulation game used to further test the students' knowledge. However, REAL is concerned with domain knowledge rather than enhancing memory and does not contain any notion of a shared event within a virtual or other world.

The work on empathic agents (e.g. [24, 27] and listening agents [3, 21] may be relevant to enhance the benefits of an RA as they have been found to deliver improved interactions. Companion agents could also potentially assist the user to reflect. For example, the "How Was Your Day" (HWYD) agent [8] is an embodied conversational agent that engages in social conversation, in contrast to personal assistant agents designed to facilitate task completion. HWYD was more than a chatterbot, though, because it had some level of understanding of the user's

utterances and was thereby able to offer advice, support or comfort after a hard day's work. Our RA was not intelligent in this respect, although its goals were nonetheless to aid learning. Also, because we are interested in enhancing and testing episodic and biographical memory of a VW event, the RA's role is to share and assist the reflection process without influencing or commenting on what is remembered.

## 1.2 Research questions and hypotheses

Drawing together educational psychology, as well as sociocognitive developmental studies examining the scaffolding of autobiographical memory, the present study aimed to investigate (1) whether reminiscing about a virtual event would improve semantic knowledge learnt within that event, when compared to no reminiscing, and (2) whether highly elaborative reminiscing about the virtual event would improve memory for the semantic knowledge learnt within that event more than low elaborative reminiscing. High elaboration is considered a holistic style, with questions and details confounded, as this best reflects observable real-life patterns of maternal scaffolding. In the present study, therefore, these components were separated in order to investigate the effect of each on memory. Our hypotheses are:

1. Semantic memory recall will be better for participants in the "reminiscing" conditions than participants in the "no reminiscing" control condition.
2. Semantic memory recall will be better for participants in the "open-ended questioning" conditions than participants in the "closed-ended questioning" conditions.
3. Semantic memory recall will be better for participants in the "high detail" conditions than for participants in the "low detail" conditions.
4. An interaction will occur between questioning and detail, such that open-ended questioning will improve semantic memory recall more for the high detail conditions than for the low detail conditions.

The RA plays a central and scaffolding role by providing guidance and sharing in the virtual event. It then questions the participant so as to assist them to reflect upon the event. By answering the above research questions we can better understand what reminiscing conversational style would be most appropriate for an RA to use to support learning.

## 2. METHOD

### 2.1 Participants

One hundred and twenty-two first-year undergraduate psychology students responded to an online advertisement. Twelve participants were excluded from the study because they failed to complete the full experiment, six were excluded due to technical problems installing the VW software, and six were excluded because they reported writing information down during the learning phase of the study. The final sample therefore consisted of 98 participants: 79 females ( $M = 21.62$ ;  $SD = 9.85$ ) and 19 males ( $M = 19.32$ ;  $SD = 11.75$ ). Participants were aged between 18 and 42 years ( $M = 20.19$ ;  $SD = 4.31$ ), and 56% identified themselves as Australian. Neither age,  $F(4,93) = .649$ ,  $p = .794$ ; gender,  $X^2(4, N = 98) = 5.14$ ,  $p = .273$ ; or cultural identity differed by condition,  $X^2(4, N = 91) = 0.95$ ,  $p = .917$ . Informed consent was obtained prior to the study. Participants were randomly allocated to conditions using an automated script. The exclusion of 24 participants resulted in slight discrepancies in numbers for each condition as given in Table 1.

**Table 1. Condition allocation and demographics**

Condition	Number	Gender (M : F)	Age M (SD)
Highly elaborative	20	5 : 15	18.60 (0.68)
Open-low	20	2 : 18	20.05 (3.52)
Closed-high	20	3 : 17	20.25 (4.28)
Low elaborative	19	4 : 15	22.53 (6.96)
Control	19	6 : 13	19.67 (3.11)

## 2.2 Experimental Design

The study was designed to investigate the impact of reminiscing talk style on students' memory of a VW. As shown in Table 2, a 2 x2 factorial between-subjects design was used to assess the impact on student learning of two independent variables: 1) Questioning style during reminiscing (open-ended or closed-ended) and 2) Provision of detail during reminiscing (high or low).

**Table 2. Reminiscing conditions**

	High detail	Low detail
Open-ended questions	Condition one ("highly elaborative")	Condition two ("open-low")
Closed-ended questions	Condition three ("closed-high")	Condition four ("low elaborative")

In addition to the four reminiscing conditions (highly elaborative, open-low, closed-high, and low elaborative), a fifth condition (control) was also included which consisted of no reflective reminiscing. The dependent variable, semantic memory recall of the VW, was measured using an online test of free recall and multiple choice items.

## 2.3 Materials

A 3D VW developed with "Unity3D" (<http://unity3d.com>) was customized for the purposes of the present study using L3DT (<http://www.bundysoft.com/L3DT>) to include just the "Omosan" village to be used for the learning phase and a virtual research laboratory resembling a metropolitan environment geographically separate from the virtual village to be used for the "reminiscing phase" of the experiment. The virtual characters and dialogues created for the study are discussed further below. The participant's "character" was designed to be of gender-neutral appearance in order to minimise gender bias (see "you" in Fig. 1).

## 2.4 Procedure

Following ethics approval, a pilot study (n=15) was conducted on friends and family of the author. All participants rated the experimental instructions and ease of navigation as being either good or very good, and no floor or ceiling effects were found for any question in the memory test. Responses to each question were in the direction we expected. The distractors in each question were therefore discriminating participants based on condition.

The study consisted of three phases and two separate time sessions. The first session was to learn as much as possible about Omosan culture through direct interaction with a number of virtual characters within the VW (learning phase). Participants would then reminisce about what they had learnt with a Reminiscing Agent (reminiscing phase). The second session involved trying to recall this information in a memory test 2 to 3 days later (test phase). Participants were also provided with experimental instructions (i.e., how to interact with objects and navigate using keyboard controls). Free navigation was deemed important to the experiment in order to better recreate an autobiographical memory, and to allow greater user control and interest in the task. Lastly, the use of pen and paper to record

dialogue within the experiment was prohibited, and participants were advised to only complete the experiment where they would not be distracted. Formal consent was obtained by clicking a checkbox and access to Omosa was achieved by clicking on a website link at the bottom of the webpage.

### 2.4.1 Learning Phase – The Village.

Participants experienced the Omosa world from the first-person perspective. A "task menu" displayed the participant's goals (e.g., "Talk to the hunter"), and progression to the reminiscing phase could not be achieved until all tasks had been completed. In the village the participant is escorted by Alan, the researcher to meet 3 virtual characters: Jennifer the resident ecologist, Yana the female hunter, Henran the Village Elder. Alan is later the RA. This means that the events/facts reflected upon and recalled relate to shared memories with RA.

The characters participate in a scripted dialogue about the fictitious cultural practices of the indigenous "Omosan people". The dialogue is related to 1) Hunting and gathering practices (e.g., "To make hunting more effective, we dip our arrow heads into *Elm sap, a sticky poisonous substance found within Elm plants*"); 2) Belief systems (e.g., "We believe there exist medicinal powers within the flesh of animals. The loin symbolises strength and so is given to our sons and the belly symbolises fertility and so is given to our daughters"); 3) Social roles (e.g., "The village elder provides guidance on spiritual and community matters, and acts as the spokesperson at ceremonial gatherings"); and 4) Social customs and rituals (e.g., "When the moon is fullest in the sky, men and women partake in ceremonial fire dancing"). Dialogue appeared on-screen as text.

Although the event had no time limit, each dialogue script between the participant and virtual characters was only displayed once, was identical for all participants, and was initiated by clicking on the virtual characters with their mouse. Dialogue continued until each of the characters had been conversed with, and all dialogue script had been exhausted. To prevent memory effects from occurring due to the ordering of information, questions and answers were presented in the same order, though the order in which the participant met a character could be controlled by the participant.

Although episodic information was not directly assessable, participants nonetheless explored environmental locations within Omosa (e.g., crop plantation) before progressing to the reminiscing phase. Such exploration was designed to increase the degree to which participants felt immersed in the VW experience, and to activate episodic memory processes. For example, it was expected that participants would recall the types of crops Omosans eat not just because they had been directed to them but because they had seen them (virtually) in person.

After participants had conversed with each of the conversational characters, they were automatically transported from the VW of Omosa to a virtual research laboratory to participate in the reminiscing phase of the experiment.

### 2.4.2 Reminiscing Phase – the Research Laboratory.

Participants could freely navigate the research laboratory in the same way as Omosa, but now they had only one objective: to talk to the RA. Participants began the reminiscing phase by clicking on the character with their mouse. Eleven free recall questions about Omosan culture were generated for the reminiscing phase of the experiment. The 11 dialogue questions focused on a specific aspect of Omosan culture and were presented in the same order

for all participants, but were structured differently, according to the experimental condition to which they had been randomly allocated (*high elaborative, open-low, closed-high, low elaborative*). See Table 3 for an example of a reminiscing question across conditions. Participants were required to respond to each question by typing into a free text input box (see Fig. 1). Questions could not be skipped without typing in a response, and in no situation did the reminiscing character confirm or refute participants' answers. To ensure the voracity of the reminiscing manipulation, participant responses were recorded in Excel; though they were not analysed further.

Participants in the *control* condition did not participate in reminiscing. Instead, they were provided with the following statement: “Welcome back! I don’t have time to discuss Omosa right now I’m sorry! Feel free to wander around the Research Lab before leaving though!” Although it was not possible to keep participants in the *control* condition for the same duration of time as those in the reminiscing conditions, it was nonetheless important to expose the *control* condition to the same environment as the reminiscing conditions, in order to help ensure that any memory effects were not the result of environmental differences between conditions.

**Table 3. Reminiscing questions across treatment conditions.**

	High Detail	Low Detail
Open-ended Questioning	Condition 1 (Highly elaborative) “What did the village elder at the back of the village say about how society was hierarchically structured?”	Condition 2 (Open-low) “How was society hierarchically structured?”
Closed-ended Questioning	Condition 3 (Closed-high) “Did the village elder at the back of the village say how society was hierarchically structured?”	Condition 4 (Low elaborative) “Is there a social hierarchy in Omosa?”



**Figure 1. Screenshot of “reminiscing phase”**

The learning and reminiscing phase lasted approximately 20 minutes in total, and participants exited the session by clicking on an “exit” button displayed on-screen.

### 2.4.3 Testing Phase

Invitations to complete the online memory test were emailed to participants 48 hours after they had completed the reminiscing phase of the study and participants had 48 hours to complete the memory test. The test phase lasted approximately 10 minutes.

A twenty-question test was created using the online survey creation software “Qualtrics”. Thirteen questions comprised the memory test, four questions related to user experience, and three questions were demographic (i.e., gender, age and cultural

identity). The thirteen-question memory test consisted of one free recall question and 12 multiple-choice questions with four items each. The free recall question asked participants: to recall in as much detail as possible everything they could remember about Omosan culture. A free text input box was provided where participants could type their answers. In terms of scoring, a coding scheme was adopted [23] whereby independent units of dialogue from the learning phase were coded. A total of 87 possible units of information from the event were generated, and recall was scored according to this coding scheme, with one point given for each unit of correctly recalled information. Consider, for example: “The Omosans bartered weaponry, jewellery, and a smoke-able herb called Marmal”. If a participant simply recalled that Omosans “bartered”, they received one point. If they identified that Omosans bartered “weaponry”, “jewellery” and a “smoke-able herb”, they received an additional three points (one point for each correctly recalled item). If a participant correctly identified the name of the smoke-able herb, “Marmal”, an additional point was given. Thus, in this case, 5 points would be allocated in total.

In order to decrease experimenter bias, only correctly spelt names of novel items (e.g. “Marmal”) were counted as a unit of information. Any information that was incorrectly recalled was counted as a distortion (i.e., incorrectly describing or mixing up information discussed within the learning phase) or an intrusion (i.e., recalling information that was never discussed in the learning phase) [25]. Furthermore, only semantic information presented as dialogue in the learning phase was scored; episodic information, such as the appearance of characters, or the number of huts inside the village, was not relevant to the study aims, and was thus omitted from analyses. Twenty percent of responses were read and scored by an independent coder who was blind to the experimental conditions and study aims. Cohen’s kappa coefficient was 0.73, indicating good inter-rater reliability.

The 12 remaining multiple choice questions tested participants’ knowledge of dialogue learnt within the learning phase. Questions related to hunting and gathering practices (e.g. *How did the Omosans make hunting more effective?*), belief systems (e.g. *What medicinal power did loin meat have?*), social roles (*What did the village elder say his role was within Omosa?*), and social customs and rituals (*When did the Omosans partake in ceremonial fire dancing?*). Importantly, 50% of the multiple-choice questions tested for semantic information that had been discussed by the RA in both the learning phase and the reminiscing phase; whilst the remaining 50% of questions tested for semantic information that had been encountered by participants in the learning phase only. Although differences in reminiscing style across experimental conditions made it impossible to reproduce questions exactly from the reminiscing phase to the memory test, it has nonetheless been found that repeated exposure to information, to any degree, improves memory recall more than does single exposure [35]. Internal consistency reliability was not assessed for the multiple-choice component of the test because items were not homogenous. Each question was worth one point with a maximum score of 12.

After the memory test, several checks of the experimental procedure were made. First, participants were asked to rate the experimental instructions, ease of navigation, and enjoyment of the task on a 5-point Likert scale (“very poor” (1) to “very good” (5)). Second, participants were asked to indicate whether or not they had copied any dialogue down during the learning phase. Six participants indicated yes and so were excluded from further analysis. Third, to determine whether purposeful rehearsal improved memory recall, participants were asked whether they

had mentally rehearsed any information learnt within Omosa during the 2 to 3 day period before the memory test using a 5-item multiple-choice scale ("No, not at all" (1) to "Yes, most of the time" (5). Finally, the frequency with which participants played video or online games was determined using a 5-item multiple-choice scale ranging from "Never" (1) to "Daily" (5).

### 3. RESULTS

The design of the current study conformed to a 2 (question style: open-ended or close-ended) x 2 (provision of detail: high or low) design with an extra condition (control). A conservative approach was taken with respect to power by applying a single factor analysis of variance (ANOVA) structure on the data. This allowed for conventional main effects and interaction tests to be conducted on the data conforming to this 2 x 2 structure, but also allowed us to compare the reminiscing conditions with the control group, and allowed any simple effects of interest to be pursued. All results were analysed while controlling for 'prior game usage', 'navigation in omosa', and 'enjoyment within omosa' that they may have an influence on the dependent variable. Due to the conservatism of the approach, alpha was set separately at  $p = .05$  for the main effects and interaction analysis, and a Bonferroni adjusted alpha of  $p = .025$  was set for the reminiscing to control group comparison, and the simple effect comparisons. An overall F-test was not considered for the 2 x 2 structure since we utilised a priori contrasts. The dependent variable, memory recall of the virtual world, was measured by an online test including free recall and multiple choice questioning.

#### 3.1 Preliminary analyses

A preliminary analyses of participants' prior experience with video games and experience of the current activity were conducted in order to determine any important covariates. Twenty-nine participants (29.90%) reported having no prior history of online or video game usage. Of the participants who did play games, 39 (40.21%) did so rarely, 15 (15.46%) played on a monthly basis, eight (8.25%) played on a weekly basis, and six (6.19%) played games daily. Importantly, game experience did not vary by condition and did not relate to higher or lower scores on the memory test,  $F(4,93) = .64, p = .633, d = .20$ . Next, participants' ratings for clarity of instructions, ease of navigation and enjoyment of the VW were assessed. Overall, the majority of participants rated these factors as being either "good" or "very good". Less than 10% of participants rated their experience as being poor or very poor (see Table 4). These participants did not vary by condition for instructions;  $F(4,93) = .61, p = .659, d = .19$ ; navigation,  $F(4,93) = 1.08, p = .372, d = .33$ ; nor enjoyment,  $F(4,93) = .77, p = .548, d = .24$ , and did not display higher scores on the memory test.

**Table 4. User ratings for clarity of instructions, ease of navigation and enjoyment.**

	very poor	poor	neutral	Good	very good
Instructions	1.03%	-	10.31%	52.58%	36.08%
Navigation	1.03%	7.22%	19.59%	45.36%	26.80%
Enjoyment	2.06%	4.12%	37.11%	39.18%	17.53%

Participants' learning behaviour between the VW and the memory test was assessed. Seventy participants (72.16%) did not mentally rehearse any of the dialogue during the 2 to 3 day period before the memory test. Twenty participants (20.62%) rehearsed once, and seven participants (7.22%) rehearsed twice or more. These participants did not vary by condition or display higher scores on the memory test,  $F(4,93) = .99, p = .419, d = .30$ . Finally, the time

taken to complete the VW event was recorded. Participants in the open-ended questioning conditions (*highly elaborative* and *open-low*) ( $M = 18.48, SD = 6.30$ ) spent significantly longer in the VW event than participants in the closed-ended questioning conditions ( $M = 15.92, SD = 5.01$ ) (*closed-high* and *low elaborative*),  $F(4,93) = 4.35, p = .04, d = .54$ . No other differences in conditions were found,  $F_s(4,93) < 2.48, p_s > .119$ .

#### 3.2 Free recall

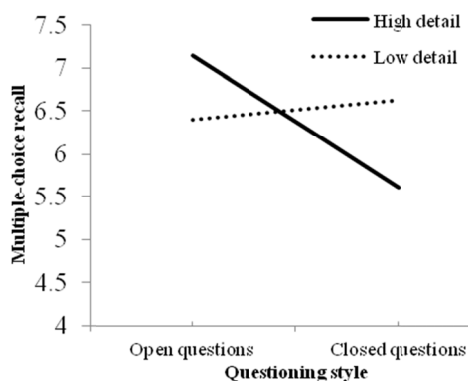
To assess the impact of reminiscing on free recall, orthogonal contrasts were used. First, free recall in the control group was compared to free recall in all reminiscing groups. No significant differences were observed,  $F(4,93) = 1.85, p = .177, d = .27$ . However, when simple effects were utilised, an independent samples t-test revealed a strong and significant difference between the *highly elaborative* condition ( $M = 18.90, SD = 9.12$ ) and the *low elaborative* condition ( $M = 12.05, SD = 6.88$ ),  $t(37) = 2.64, p = .012, d = .87$ ; as well as between the *highly elaborative* condition ( $M = 18.90, SD = 9.12$ ) and the *control* group ( $M = 12.42, SD = 8.06$ ) for the free recall component of the memory test,  $t(37) = 2.35, p = .024, d = .77$ .

Second, within the reminiscing groups, the main effects and interactions for questioning type and provision of detail were examined. These tests revealed a significant main effect for questioning type,  $F(4,93) = 4.62, p = .034, d = .57$ , whereby participants in the open-ended questioning conditions (*highly elaborative* and *open-low*) ( $M = 17.18, SD = 9.54$ ) freely recalled more correct information than did those in the closed-ended questioning conditions (*closed-high* and *low elaborative*) ( $M = 13.35, SD = 7.79$ ). There was no significant main effect for detail,  $F(4,93) = 2.59, p = .111, d = .36$ , and no significant interaction between questioning type and detail,  $F(4,93) = .76, p = .385, d = .14$ .

**3.2.1. Free Recall Errors.** To complement the analyses of correct free recall, the number of distortions (i.e., incorrect describing or mixing up of information) and intrusions (i.e., recalling information never part of Omosa) made between conditions was also assessed. Interestingly, participants in the closed-ended questioning conditions (*closed-high* and *low elaborative*) made significantly more free recall memory distortions ( $M = 0.89, SD = 1.08$ ), on average, than did participants in the open-ended questioning conditions (*high elaborative* and *open-low*) ( $M = 0.45, SD = 0.72$ ),  $F(4,93) = 4.21, p = .043, d = .53$ . There was no main effect for provision of detail, and no significant interaction,  $F_s(4,93) < .54, p_s > .466$ . Likewise, no significant differences between conditions were found in the total number of intrusions made,  $F_s(4,93) < .37, p_s > .542$ .

#### 3.3 Multiple-choice recognition

To assess the impact of reminiscing condition on multiple-choice recognition scores, the same orthogonal contrasts were used as for the free recall. There was no significant difference in correct multiple choice recognition between the reminiscing and control condition,  $F(4,93) = .08, p = .780, d = .06$ . When examining the main effects and interaction effects of questioning and details within the reminiscing conditions, however, some differences between free recall and multiple-choice recognition emerged. For multiple-choice recognition, neither the questioning,  $F(4,93) = 2.13, p = .148, d = .30$ , nor detail main effects were significant,  $F(4,93) = 0.13, p = 0.719, d = 0.07$ . There was nonetheless a significant interaction between questioning type and detail,  $F(4,93) = 5.31, p = 0.024, d = 0.63$ . Specifically, the *highly elaborative* condition ( $M = 7.15, SD = 2.41$ ) recalled more multiple-choice items, on average, than the *open-low* condition ( $M = 6.40$ ,



**Figure 2. Interaction between questioning and detail for multiple-choice questions.**

$SD=1.98$ ); whereas the *low elaborative* condition ( $M = 6.63$ ,  $SD = 2.09$ ) recalled more multiple-choice items, on average, than for the *closed-low* condition ( $M=5.60$ ,  $SD= 2.11$ ). See Fig.2.

In order to further test for differences in multiple-choice recognition between those aspects of the Omosa VW which participants had been exposed to just once (i.e. elements that were present in the event itself but not during reminiscing) and aspects to which participants had been exposed to twice (i.e. elements that the reminiscing character addressed directly), the multiple-choice test was designed with six “single-exposure” questions and six “double-exposure” questions. Interestingly, no significant main effects or interactions were observed for the double-exposure questions,  $F_s(4,93) < 1.93$ ,  $p_s > .169$ , nor for the main effects for the single-exposure questions,  $F_s(4,93) < .94$ ,  $p_s > .335$ . Nonetheless, there was a significant interaction,  $F(4,93) = 8.68$ ,  $p = .004$ ,  $d = .83$ : participants in the *highly elaborative* condition ( $M = 3.25$ ,  $SD = 1.45$ ) recalled more single exposure questions, on average, than the *open-low* condition ( $M = 2.30$ ,  $SD = 1.34$ ); whereas the *low elaborative* condition ( $M = 2.79$ ,  $SD = 1.27$ ) recalled more single exposure questions, on average, than the *closed-high* condition ( $M = 2.20$ ,  $SD = 1.11$ ). This pattern matches the interaction pattern for multiple-choice responses more generally, thus suggesting that these differences in multiple-choice responses are driven by differences in responses to the single-exposure questions.

#### 4. DISCUSSION

This study was designed to investigate the impact of an RA’s reminiscing talk style on students’ memory of a VW. Hypothesis One predicted that semantic memory recall would be better for participants in the reminiscing conditions (*highly elaborative*, *open-low*, *closed-high* and *low elaborative*) than for participants in the non-reminiscing condition (*control*). However, no significant differences were found for either the free recall or multiple-choice components of the memory test. As we utilised a priori testing, it is possible that combining the reminiscing conditions into one analysis was inappropriate. However, when simple effects were examined, participants in the *highly elaborative* condition recalled significantly more information in free recall than participants in the *low elaborative* condition.

This finding is consistent with a study [29] which revealed that mothers exposed to highly elaborative reminiscing training across a one year period had children with significantly better event memory recall than children of mothers without such training.

This finding also supports research [9, 12] which reports better autobiographical memory recollection in children of highly elaborative mothers, compared to low elaborative mothers. Furthermore, participants in the *highly elaborative* condition also showed superior free recall memory when compared to the *control* group. As participants in the *control* group did not receive a repeated exposure to the dialogue during the reminiscing phase questions, this finding supports a multitude of studies that have reported benefits for reflection on learning and memory [10, 26].

Hypothesis Two predicted that the open-ended questioning conditions (*highly elaborative* and *open-low*) would enhance memory recall better than the closed-ended questioning conditions (*closed-high* and *low elaborative*). Although this finding was not observed for the multiple-choice items, those in the *highly elaborative* and *open-low* conditions performed significantly better in the free recall component of the memory test. This finding is supported by a multitude of research reporting benefits for using open-ended during classroom conversation [20, 30].

Hypothesis Three in turn predicted that provision of high detail (*highly elaborative* and *closed-high*) would enhance memory recall more than low detail (*open-low* and *low elaborative*). Contrary to this prediction, no significant differences were observed for either the free recall or multiple-choice items. While no studies to date have examined the provision of detail independently of open-ended questioning, a multitude of studies nonetheless highlight the important role of highly detailed, open-ended reminiscing (that is, high-elaborative reminiscing) for children’s recall (e.g. [9, 12, 29]). Our findings suggest that the provision of detail alone is not sufficient to improve memory.

Finally, Hypothesis Four predicted an interaction between questioning type and detail such that open-ended questioning would improve memory recall more for the high detail condition than for the low detail condition. In partial support for this hypothesis, this interaction effect was observed for multiple-choice recognition (but not free recall). Drawing together the findings from Hypothesis Three and Hypothesis Four, this finding suggests that details enhance students’ recall of an experiential virtual event only when other memory processes are strongly supported: that is, when open-ended questions are asked during the reflective encoding phase, and when memory prompts in the form of multiple-choice questions are provided during recall.

It appears that the *nature* of the reminiscing and the scaffolding is critical. Interestingly, the *closed-high* condition (closed-ended questioning and high detail) was not sufficient to promote deep encoding. Thus, there appears to be a critical difference between reminiscing with younger children, who developmentally may need this extra scaffolding (i.e., details), and young adults, who are more able to direct their own learning and elaborate on missing information more successfully. Indeed, in adults, extra detail in the closed-ended questioning condition appears to have restrained the degree to which participants could elaborate.

That open-ended questions during reflective reminiscing promoted better free recall, irrespective of the provision of detail, supports various studies that have also found educational benefits for using open-ended questioning in classroom settings. Specifically, not only do open-ended questions lead to greater topic engagement and motivation [33], they also encourage deeper metacognitive reflection and higher-order thinking than closed-ended questions allow [30]. Compared to closed-ended questions, then, open-ended questions help consolidate memories more strongly in the brain, and result in better memory performance.

Interestingly, closed-ended questioning (*closed-high* and *low elaborative*) showed significantly more distortions in free recall than open-ended questioning (*highly elaborative* and *open-low*). In addition to facilitating higher-order thinking and elaborative self-generated responding, it appears that open-ended questions assist with the encoding and consolidation of more *accurate* memory representations, compared to closed-ended questions. As closed-ended questions discourage elaborate thinking and responding, more distortions in memory are probable since the memory representation is weaker and less stable to begin with.

It is interesting that an advantage for open-ended questioning during reflective reminiscing was not observed for the multiple-choice component of the memory test. As multiple-choice involves recognition memory, the results seem to suggest that open-ended questions may not be necessary in this context. This means it is only in the absence of other memory cues (i.e., as occurs during free recall) that the stronger encoding driven by open-ended questioning is needed. When additional memory cues are provided, in the form of multiple-choice question and response sets, then even weaker memory traces can be recalled.

While the main effect for questioning was significant for free recall, there was no difference for the main effect for details for either free recall or multiple-choice. This may be because reminiscing occurred immediately following the event. In other autobiographical event memory research, the reminiscing phase occurs at a time-point later than the original event. Second, it appears that open-ended questions are more important for learning and memory than the provision of high detail. In child developmental studies, the two are typically confounded (i.e., it comprises both open-ended questions and high detail).

Nonetheless, it must be remembered that a significant interaction was observed between questioning type and detail for the multiple-choice items. Specifically, the inclusion of “detail” to open-ended questioning (i.e., *highly elaborative*) provided an additional “booster” effect to recognition memory performance, compared to just open-ended questioning alone. The inclusion of these extra details during reminiscing might have provided participants with additional memory cues not available in the *open-low* condition. Not only might these cues help the reconsolidation of existing event memory representations, they could also permit the formation of new memory linkages previously forgotten, or unattended to, in the learning phase. These additional memory cues would be expected to create more stability in the participant’s memory representation, and as such, improve memory recall on subsequent tests.

The opposite effect was found for the closed-ended questioning conditions, however. That is, participants in the *low elaborative* condition showed better memory recall than the *closed-high* condition, even though the *closed-high* condition contained more detail during reminiscing. Although this finding appears paradoxical, the inclusion of extra detail during reminiscing may only be beneficial for learning if the opportunity to encode deeply (i.e., through open-ended responding) is also available. Without deep encoding through open-ended responding, the provision of high detail appears to be damaging to memory.

Consistent with this idea, when recall was divided between single exposure questions (i.e., questions drawn from the learning phase only) and double exposure questions (i.e. questions drawn from both the learning phase and reminiscing phase), an interaction was observed for single exposure questions only. That is, participants in the *highly elaborative* condition recalled the answers to more

multiple-choice single exposure questions than those in the *open-low* condition; whereas participants in the *low elaborative* condition recalled more than those in the *closed-high* condition. Again, the provision of high detail in open-ended questioning (i.e., *highly elaborative*) appears to be beneficial for learning. This suggests that these participants were able to build up a stronger schema of the entire VW event in order to remember these single exposure questions. Indeed, it appears that participants in the high detail closed-ended condition (*closed-high*) had a weaker schema of the VW event, and so were more inclined to extract information from the double exposure questions which were explicitly told to them (i.e., their memory was weaker for the non-discussed parts of the event). That participants performed equally well in the double exposure multiple-choice questions suggests that high elaboration benefits memory most when less encoding opportunities (i.e., the single exposure questions) have initially been provided.

## 5. CONCLUSIONS AND FUTURE WORK

Notwithstanding these findings, however, limitations and directions for future research can be identified. First, a significant increase in time to complete the responses was found for the open-ended questioning conditions (*highly elaborative* and *open-low*) compared to the closed-ended questioning conditions (*closed-high* and *closed-low*). Given the confound between time and questioning, it is unclear whether a similar benefit for questioning would emerge when open-ended questions are used but greater time pressure is exerted. As this preliminary, investigative experiment only required participants to read and remember text-based dialogue during interaction with an IVA and as there was no specific problem or challenge that needed to be solved (except for remembering the dialogue for future recall), it is possible that our study was not challenging enough. Future studies should utilise a more complicated, problem-based learning task with greater unpredictability of goal attainment. Third, it would be worthwhile to temporally space the learning phase and reminiscing phase apart (e.g., at 2 days, at 4 days, at 1 week, etc), instead of having them in immediate succession. Although beyond the scope of the present study, future studies should also analyse responses generated in the reminiscing phase, as well as the memory test phase. Finally, we focused on reflection on *content* only (i.e., how well students remembered information learnt within a VW event after exposure to different reflective reminiscing styles). However, the experimental paradigm used in the present study might also be utilised to consider reflection on the learning *process* itself.

This study offers preliminary research into the effectiveness of highly elaborative reminiscing on semantic memory recall of a VW. This has implications for the design of embodied conversational agents beyond the educational context. Our findings suggest what conversational style might be more appropriate for companion agents involved in improving and retaining day to day memory skills. This style involves highly elaborative reminiscing, comprised of both open-ended questions and a provision of high detail, and is superior to low elaborative reminiscing (i.e., closed-ended questions and low detail), and no reminiscing, during free recall memory tasks. Open-ended questioning conditions also reported fewer memory errors in free recall than did participants in the closed-ended questioning conditions. Our study also suggests that the provision of detail is damaging to memory if the opportunity to elaborate using open-ended responding is unavailable.

Interestingly, these findings emerged even when the reminiscing partner was a virtual character insensitive to the students' responses. With a more intelligent agent with social ability, listening and collaborative skills and possibly memories of its own, the benefits of utilizing an RA to enhance student memory produced in this study is likely to be magnified.

## 6. REFERENCES

- [1] Aylett, R., Vannini, N., Andre, E., Paiva, A., Enz, S., & Hall, L.E. (2009) But that was in another country: agents and intercultural empathy. Proc. AAMAS09. Budapest, Hungary, 329-336.
- [2] Bai, X., Black, J. B., Vitale, J. (2007). REAL: An agent-based learning environment. Paper presented at the Agent-Based Systems for Human Learning Conference, Hawaii.
- [3] Bevacqua, E., Mancini, M., and Pelachaud, C. 2008. A Listening Agent Exhibiting Variable Behaviour. In Proc. IVA'08 Tokyo, Japan, Sept01 - 03, 2008. Prendering, Lester, & Ishizuka (eds) LNCS 5208. Springer-Verlag, Berlin, Heidelberg, 262-269.
- [4] Bickmore, T. W., Schulman, D., & Yin, L. (2009) Engagement vs. Deceit: Virtual Humans with Human Autobiographies. IVA 2009. 6-19.
- [5] Brom, C., & Lukavský, J. (2009) Towards More Human-Like Episodic Memory for More Human-Like Agents. In: Ruttkay, Zs. et al. (Eds): IVA 2009, LNAI 5773, 484-485.
- [6] Burkert, O., Brom, C., Kadlec, R., & Lukavský, J. (2010) Timing in Episodic Memory: Virtual Characters in Action. In Proc AISB 2010, Lim & Wan (Eds.), De Montfort Uni., Leicester, UK
- [7] Campos, J. and Paiva, A. (2010) MAY: my memories are yours. In Proc. IVA'10, Allbeck, Badler, Bickmore, Pelachaud & Safonova (Eds.). Springer-Verlag, Berlin, Heidelberg, 406-412.
- [8] Cavazza, M., Santos de la Cámara, R., Turunen, M. and the COMPANIONS Consortium (2010) How Was Your Day? a Companion ECA. In Proc.AAMAS 2010, 10-14 May 2010, Toronto, Canada.
- [9] Cleveland, E. S., & Reese, E. (2005). Maternal structure and autonomy support in conversations about the past: Contributions to children's autobiographical memory. *Developmental Psychology*, 41, 376-388.
- [10] Craik, F. I. M. 2002. Levels of processing: past, present and future? *Memory*, 10, 305-318.
- [11] Dias, J., Ho, W. C., Vogt, T., Beeckman, N., Paiva, A., & André, E. (2007) I know what you did last summer: Autobiographic memory in synthetic characters. Proc. ACII 2009 (pp. 606--617). Lisbon, Portugal
- [12] Fivush, R., Reese, E., & Haden, C.A. (2006). Elaborating on elaborations: Role of maternal reminiscing style in cognitive & socioemotional development. *Child Devel.*, 77:1568-88.
- [13] Gomes PF, Martinho C, Paiva A. 2011. NPC memory affecting emotional reactions and path planning. *Game AI*.
- [14] Ho, W. C., Dautenhahn, K., & Nehaniv, C. L., (2008) Computational memory architectures for autobiographical agents interacting in a complex virtual environment: a working model, *Connection Science*, 20:1, 21-65
- [15] Hoscheidt, S. M., Nadela, L., Payne, J., & Ryana, L. (2010). Hippocampal activation during retrieval of spatial context from episodic and semantic memory. *Behavioural Brain Research*, 212, 121 - 132.
- [16] Kasap, Z., Ben Moussa, M., Chaudhuri, P., & Magnenat-Thalmann, N. 2009. Making Them Remember - Emotional Virtual Characters with Memory. *Computer Graphics and Applications, IEEE 29(2): 20-29.*
- [17] Kolb, D. A. (1984) *Experiential learning: Experience as the source of learning and development*. NJ: Prentice-Hall.
- [18] Lee, Y., Kinzie, M. B., & Whittaker, J. V. (2012). Impact of online support for teachers' open-ended questioning in pre-k science Activities. *Teaching & Teacher Educ.*, 28:568-577.
- [19] Lim, M. Y., Aylett, R., Ho, W. C., Enz, S., & Vargas, P. (2009) "A Socially-Aware Memory for Companion Agents" In: Zs Ruttkay et al. (Eds): IVA 2009, LNAI 5773, 20-26.
- [20] Lim, M.Y. Ho, W. C., Aylett, R. 2011. Human-like Memory Retrieval Mechanisms for Social Companions, Proc. of 10th AAMAS 2011, Yolum, Tumer, Stone and Sonenberg (eds.), May, 2-6, 2011, Taipei, Taiwan
- [21] Maatman, R. M. and Jonathan Gratch and Stacy Marsella, 2005. Natural behavior of a listening agent, *Lecture Notes in Computer Science*, Springer-Verlag, London, UK
- [22] Mathews S., Andrews, L., & Luck, E. (2012). Developing a Second Life virtual field trip for university students: an action research approach. *Educational Research*, 54, 17-38.
- [23] McCabe, A., Peterson, C., & Connors, D. M. (2006). Attachment security and narrative elaboration. *International Journal of Behavioural Development*, 30, 8-19.
- [24] McQuiggan, S. W., and Lester, J. C. 2006. Learning empathy: a data-driven framework for modeling empathetic companion agents, Proc. AAMAS'06, May 08-12, 2006, Hakodate, Japan
- [25] McGuigan, F., & Salmon, K. (2004). The time to talk: The influence of the timing of adult-child talk on children's event memory. *Child Development*, 75, 669-686.
- [26] Moon, J. (1999). *Reflection in learning and professional development: Theory and practice*. London Sterling, VA: Kogan Page.
- [27] Niewiadomski, R., Ochs, M., and Pelachaud, C. 2008. Expressions of Empathy in ECAs, Proc. IVA'08, 37 - 44
- [28] RADAR (2007) Carnegie Mellon University [http://www.ri.cmu.edu/research\\_project\\_detail.html?project\\_id=585&menu\\_id=261](http://www.ri.cmu.edu/research_project_detail.html?project_id=585&menu_id=261) accessed 12-October 2012
- [29] Reese, E., & Newcombe, R. (2007). Training mothers in elaborative reminiscing enhances children's autobiographical memory and narrative. *Child Development*, 78, 1153-1170.
- [30] Roth, W. M. (1996). Teacher questioning in an open-inquiry learning environment: interactions of context, content, and student responses. *Journal of Research in Science Teaching*, 33(7), 709-736.
- [31] Schoen, D. (1983). *The Reflective Practitioner: How professionals think in action*. London: Temple Smith.
- [32] Sims, R. & Bovard, B. (2004). Interacting with online learners: How new elaborations of online presence can foster critical thinking and reflection. In R. Atkinson, C. Retrieved from <http://www.ascilite.org.au/conferences/perth04/procs/pdf/sims.pdf>.
- [33] Smith, H., & Higgins, S. (2006). Opening classroom interaction: the importance of feedback. *Cambridge Jnl of Ed.*, 36, 485-502.
- [34] Tecuci, D and Porter, B. 2007. A generic memory module for events, in Proc. 20th Florida AI Research Society Conference (FLAIRS 20), Key West, FL.
- [35] Toppino, T. C., & Bloom, L. C. (2002). The spacing effect, free recall, and two-process theory: A closer look. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 28(3), 437-444.
- [36] Vargas, P. A., Ho, W. C, Lim, M. Y., Enz, S., Fernaeus, Y., & Aylett, R. (2009) To Forget or Not to Forget: Towards a Roboethical Memory Control. In: Proc. AISB'2009, Herriot Watt University, Edinburgh, Scotland.