Can virtual environments enhance the learning of historical chronology?

Nigel Foreman · Stephen Boyd-Davis · Magnus Moar · Liliya Korallo · Emma Chappell

Abstract Historical time and chronological sequence are usually conveyed to pupils via the presentation of semantic information on printed worksheets, events being rote-memorised according to date. We explored the use of virtual environments in which successive historical events were depicted as ''places'' in time–space, encountered sequentially in a fly-through. Testing was via ''Which came first, X or Y?'' questions and picture-ordering. University undergraduates experiencing the history of an imaginary planet performed better after a VE than after viewing a ''washing line'' of sequential images, or captions alone, especially for items in intermediate list positions. However, secondary children 11–14 years remembered no more about successive events in feudal England when they were presented virtually compared with either paper picture or 2-D computer graphic conditions. Primary children 7–9 years learned more about historical sequence after studying a series of paper images, compared with either VE or computer graphic conditions, remembering more in early/intermediate list positions. Reasons for the discrepant results are discussed and future possible uses of VEs in the teaching of chronology assessed.

Keywords Chronology teaching · Virtual environments · Primacy and recency effects

Introduction

Educational research has shown time to be an extremely challenging concept for children to grasp, mainly due to its abstract nature and the absence of immediate representative

N. Foreman (✉) · L. Korallo · E. Chappell
Psychology, School of Health and Social Sciences, Middlesex University, Enfield Campus,
Queensway, Enfield EN3 4SF, UK
e-mail: n.foreman@mdx.ac.uk

S. Boyd-Davis · M. Moar
Lansdown Centre, School of Arts, Middlesex University, Queensway, Enfield, UK

sensory information (Hodkinson 1995; Haydn 1999) so that children find difficulty in placing events in time-ordered sequence (Partington 1980). The time conception relates not just to clock time but includes a multifaceted array of other elements, incorporating both calendar and seasonal along with personal and historical time (Partington 1980; Hoodless 1996; Cooper 2000). Moreover, children's ability to comprehend time does not develop in isolation but relies on an abundance of other emerging skills including numeracy, literacy, and particularly memory. It is such abilities which combine eventually to enable the child to develop proficiency in comprehending that time ''passes'' and that it is possible to quantitatively measure elapsed time (Hoodless 1996).

History learning requires an understanding of ''chronology'' (Bourdillon 1994; Hoodless 1996), defined by Smart (1996) as 'the sequencing of events/people/developments in relation to other and existing knowledge of other, already known, events/people/developments' (p. 79). Epochs must be remembered in sequence along with their relevant dates (Hoodless 1996). Masterman and Rogers (2002) point out that children have difficulties with chronology, particularly when they have to deal with past events or people that are out of their own direct experience. Once acquired, the concept of chronology eventually enables children to locate themselves in relation to historical events (Hoodless 1996).

The national curriculum (NC) for history requires that children at key stages I and II acquire awareness of time and chronology (Hoodless 1996), though the significance of such knowledge extends beyond the history curriculum per se since history can be considered an ''umbrella discipline''; it incorporates and promotes understanding of wider developments in society, music, art, religion, technology and science, enabling children to begin to understand how previous societies differed from their own and how chronologically distinct events influence one-another (see Cooper 2000).

Various strategies have been employed to attempt to enhance children's chronological thinking skills, both in non-disabled children and also those with learning difficulties who can have particular problems in this domain. Turner (1998), for example, engaged the emotional dimension by having disabled children think about ''what it would be like'' to have lived on a sailing ship; this appears to have been successful. Evidence-based approaches have often used artefacts, visual resources, and especially timelines (O'Hara and O'Hara 2001). Artefacts can encourage curiosity and interest, and leave lasting impressions (Hoodless 1996; O'Hara and O'Hara 2001; Wood and Holden 1995; Cooper 2000), though there is nothing that intrinsically ''places'' such objects in time or sequence. Representational visual materials such as photographs, paintings or posters are highly accessible, especially if children are guided in their use, and if the significance of time-relevant cues is pointed out (O'Hara and O'Hara 2001), although generally they have the advantage that they need not rely on language for their impact (Hoodless 1996; O'Hara and O'Hara 2001).

Perhaps the most common medium for teaching chronology is the timeline (O'Hara and O'Hara 2001; Smart 1996). This can use depictions of historical events, pegged across a washing line like items of clothing, or in the form of a roll chart or scroll, circular booklet, three-dimensional line, or as a vertical or horizontal two-dimensional wall display. Paper timelines are cheap and can be used to represent any type of information (Madeley 1921; Smart 1996), though to be effective they should present information in an uncluttered and unambiguous manner (Hoodless 1996). Nevertheless, for such a 2-dimensional timeline, all the content is presented contemporaneously, and thus the time element has to be imposed on the display by the viewer as they scan successive images from one end of the timeline to the other.

A better form of representation may be possible via the use of Information and Communication Technology (ICT), which has been recommended as an effective medium for history teaching (Cooper 2000; O'Hara and O'Hara 2001; Wood and Holden 1995; Yaxley

2004). It may impact in various ways on the learning process, and the organisation of historical information, being a highly motivating medium (Smart 1996; Underwood and Underwood 1990; Watson 1993; Wood and Holden 1995; see QCA 1999). Past events can be brought to life via realistic depictions (Yaxley 2004). Also, computer use is co-operative, so that the teacher can optimise learning by ensuring inclusion of all pupils, in working groups (O'Hara and O'Hara 2001).

Masterman and Rogers (2002) used interactive multimedia (IMM) as a means of ''scaffolding'' children's understanding of historical chronology, using a 2-D timeline. Wood et al. (1976) showed that children's understanding of abstract concepts can be facilitated by scaffolding, i.e. the incorporation of various kinds of learning material to teach children how to solve problems that are originally outside their competence (see also Scaife and Rogers 1996). Masterman and Rogers' (2002) IMM program was designed to teach children how to effectively select the most appropriate representations for problem solving, by providing a better understanding of temporal systems. They utilised the met-aphor of time travel along a prescribed road-map route, along which historical knowledge could be located. Information about people depicted on the route was obtainable by clicking on the corresponding icon. Dates were represented as ''years before the present time'', to avoid children's having to understand numerical dates. Masterman and Rogers (2002) received positive feedback from both teachers and children, although the superiority of the technique over other teaching media was not formally assessed.

The studies described above used semantic memory for retaining historical information. However, it could be argued that spatial memory would be more appropriate. O'Keefe and Nadel (1978) argued that the most effective spatial representations involve the construction of cognitive 'maps', which allow the understanding of spatial relationships among environmental cues and landmarks (see also Tolman 1948; Foreman and Gillett 1997, 1998). Cognitive maps enable locations to be remembered sequentially as a participant moves through and past them; they are stable, relatively permanent and of apparently unlimited storage capacity.

A recent technology that is increasingly impacting on education is virtual environment (VE) technology, which could potentially ''... change the structure of the classroom, the curriculum and the learning style of the students'' (Passig and Sharbat 2000, p. 7). There is abundant evidence that VEs can convey information regarding spatial layouts and environmental relationships, allowing route-learning, mapping, and landmark use, of a kind associated with cognitive map formation (O'Neil 1992; Richardson et al. 1999; Ruddle et al. 1997; Foreman et al. 2003). Moreover, VEs do this more effectively than simple two dimensional (2-D) representations (Ruddle et al. 1997; Stanton et al. 1996). They allow the user to explore a space equivalent to a large public building in pseudo-real time (Foreman et al. 2003) and feel a sense of presence in the depicted environment (Wilson 1997). VEs have been widely used as a spatial knowledge training medium (Foreman et al. 2003, 2005; Jacobs et al. 1997; Ruddle et al. 1997; Waller 2000). Skills acquired from the exploration of VEs have been found to be suitable for practical purposes (e.g., Bliss et al. 1997; Foster et al. 1998; Foreman et al. 2003, 2005; Tate et al. 1997). Space–time dimensions have been exploited in VEs by archaeologists to display changes which have occurred over the years on excavated sites, allowing users to take a visual journey through the historical past (Barcelo et al. 2000). Similarly, incorporating VEs into history teaching could potentially provide children with a visual fly-through of successive eras of national history, as if travelling through time-space in a time machine. Using VEs could thus transform the traditional 2-D timeline, so that the memories for historical events, names and dates could be encoded and stored within the spatial memory system, thereby enhancing retention.

Kullberg (1995, 1996) attempted to devise a dynamic 3-D timeline in which events are truly experienced sequentially, visually depicting the history of photography. The user had an overview of a 3-D landscape containing chronological lines of photographs with which they could interact by selecting specific pictures to obtain detailed information about them. While the educational impact of using such a display was not investigated (Kullberg 1995, 1996) it was argued that the utilisation of interactive environments allows seamless communication of historical information at both the micro and macro level of detail, and from a multitude of different viewpoints.

Pedley et al. (2003) investigated the use of VEs by a small number of disabled young people, who created personal timelines consisting of visual images of significant events in their lives which could be re-experienced repeatedly as a ''fly-through''. The results were promising and suggested that these individuals may have gained a better grasp of chronology as a result of their participation.

The current study further investigated this possibility using a series of specific events, experienced by passing them in a fly-through or viewing equivalent materials in printed paper form (printed worksheets, picture sequences, or a washing line timeline, materials commonly be used in classrooms) or as a sequence of 2-D computer graphics. The specific hypothesis that applied to all three studies was that more would be learned about the sequencing of events from VEs than from any 2-D sequences of successive graphics or captions, since the former incorporates 3-dimensionality and thus arguably invokes ''navigational'' spatial memory (Foreman et al. 2003; Wilson 1997). We also hypothesised that a PowerPoint graphical presentation would be more effective than a simple paper or washing line graphic presentation, due to the generally motivating effect of computer use, and that graphical presentations would improve on worksheets containing semantic descriptions alone.

A list of nine items was used in each study reported below, representing the upper extreme of the classical 7 ± 2 buffer storage capacity assumed for verbal short term memory (Miller 1956), thus allowing most information to be remembered but opportunity for improvement where a training condition proved successful. A common feature of supraspan list-learning is that while material in early and late list positions are well remembered, middle order items are most often forgotten. This is known as the serial position effect (SPE) (Atkinson and Shiffrin 1971; Glanzer and Cunitz 1966). Increasing arousal via moderate caffeine consumption tends to enhance recall from middle list positions (Barraclough and Foreman 1994). The phenomenon has been examined extensively with verbal materials but not previously with spatial or event-related materials. Should a U-shaped SPE effect appear in the present data, any benefit of VE use might be expected to appear in middle list positions, if the VE generally enhances excitement for the material. However, if VE use proves over-exciting, recall may be suppressed (cf. Broadbent 1981).

The present three studies were designed independently and conducted in parallel; designs, protocols, types of historical material and control conditions were varied to suit the different age groups (undergraduate students, 11–14 year-old, and 8–9 year-old pupils). The studies took account of participants' gender and, where possible, teacher-assessed academic ability. For the undergraduates the test materials were equally unfamiliar to all participants, since they related to the history of an imaginary planet. The 11–14 year-olds were tested with history NC-related Feudal Britain materials, to determine whether the VE could effectively supplement standard teaching. The 7–9 year-olds learned about more general eras of history, appropriate to their age. Although inter-group protocol differences precluded formal age comparisons, it was expected that undergraduate students would benefit from VE use (due to their relaxed computer use) while the younger age group may

be overexcited by the VE medium and thus remember less of the chronological material, by comparison with their corresponding 2-D conditions.

Experiment 1

Method

Participants

Participants consisted of an opportunity sample of 39 university undergraduate students, 15 males and 24 females, aged 18–22 years. Module credits were available as a reward for participation, though more than half volunteered freely. All had normal or corrected-to-normal vision.

Materials

A set of nine imaginary dates/events was created by the researcher, representing events in the history of an imaginary planet, using colour images obtained from the internet and cropped using Microsoft Paint. Each was ascribed an imaginary date, at successive 20 year intervals, and a suitable caption (see Appendix 1a). The images were pasted as successive screens in a VE (Fig. 1) thus creating a dynamic web-based timeline which the participant could move through. The VE was created using Virtools Virtual Reality software [Virtools Dev 3.0 Educational Version; see www.virtools.com] running on a Pentium 3 450 mhz computer with 128 mb RAM. The environment ran in Microsoft Explorer with a Virtools 3d plug-in. For a ''washing line'' control condition, the same materials were depicted on nine successive A4 sheets, in portrait orientation, pegged along a string which was looped

Fig. 1 View of the VE as the sequence began, in Experiment 1

across one wall of the room. For a printed text control condition, the captions and dates were printed, three per page, on three successive otherwise blank A4 sheets in portrait orientation, stapled together with a single staple in the top left hand corner.

Materials used to assess participants' memory for events were: (1) a questionnaire posing nine questions in the form ''Did X come before Y?'' requiring true/false responses, and (2) nine test pictures (as in the VE display, with captions but without dates) in A5 portrait format, scattered randomly on a desk; participants had to place them in chronological order on a wall, attaching them using blu-tack plastic adhesive.

Procedure

Participants were randomly assigned to one of three groups, with 13 in each Those in the VE group (4M, 9F) experienced their events/dates via a VE fly-through. They were seated comfortably approximately 40 cm from the screen of a PC desk-top computer, and were asked to operate a forward directional key to fly through the environment. They were asked to imagine that they were walking through the environment depicted on the screen. They could pause during the fly-through if they wished, but movement was unidirectional, backward and sideways movement keys having been disabled. After passing the last of the nine pictures, a message appeared on the screen asking ''Do you want to go again?'' to which a ''Y[es]'' response took them back to the start of the fly-through, a 15 s delay occurring before the first image came clearly into view. The total time taken to complete the fly-through, with the forward key continuously depressed, was 67 s.

Participants in the washing line group (4M, 9F) were allowed to view the same nine pictures with captions and dates in the form of the timeline, described above. The printed text group (7M, 6F) were given the printed stapled sheets, described above. All participants were instructed to study the nine events sequentially five times, familiarising themselves with the information presented and attempting to remember events, dates and order of occurrence. Upon completion of the task participants were taken to a nearby quiet room and asked to complete the questionnaire, and then place the nine A5 test pictures in a line in their correct order of occurrence. They were allowed unlimited time to do this, though all participants completed testing in approximately 4–6 min.

Results

For the initial analysis four scores were obtained for each participant. First, the number of A5 test pictures placed in their correct list positions, and second, the number of correctly answered questions (out of nine) on the questionnaire were recorded for each individual. A third score, referred to hereafter as a ''Removed1'' score, was derived from the placement data by assessing, for each picture, how far removed (i.e. the number of positions) it was from its correct list position, and obtaining a total for each participant by summing across the nine pictures. [For a particular picture which ought to be placed in position 2 but was actually placed in position 6, its Removed1 score would be $6 - 2 = 4$. A correctly placed picture would obtain a score of 0]. Finally, a ''Removed2'' score was calculated by subtracting from the total Removed1 score the score for the highest-scoring picture. Removed1 scores indicated overall placement accuracy and list organisation, while Removed2 scores minimised the effect of a high score arising from the completely inappropriate placement of one particular item, which could disrupt an otherwise well-organised sequence and lead to an unrepresentatively high score.

To examine SPEs, the number of items placed correctly in list positions 1–3, 4–6, and 7–9 were recorded separately for each participant.

Data for the first four dependent variables were well distributed, justifying the use of parametric testing. A Group × Gender 2-way analysis of variance (ANOVA) for independent groups was used followed by post hoc group comparisons using the Least Significant Differences test (with 2-tailed probabilities, unless otherwise specified). There was a significant group difference in the number of pictures placed in their correct positions, $F(2, 33) = 4.41$; $P = .02$. The VE group was significantly better than either the washing line group, $P < .02$, or the caption controls, $P < .008$. There was no significant difference between the washing line and caption groups, $P = .84$. The sexes did not differ, $F(1, 3) = 2.38$; $P = .13$, nor was there any Group × Gender interaction, $F(2, 33) = 48$; $P > .05$. For the number of questions answered correctly, an almost significant group difference was obtained, $F(2, 33) = 2.99$; $P = .064$, mean scores for the VE group being higher than either the washing line or caption control groups. Males' scores (mean: 8.13) were generally higher than females' (mean: 7.12), a result that almost reached significance, $F(1, 33) = 3.59$; $P = .067$. There was no Group × Gender interaction, $F(2, 33) = .04$; $P = .96$.

Removed1 scores (Fig. 2) showed a highly significant group difference, $F(2, 33) = 5.95$; $P < .007$, and a nearly significant gender difference favouring males, $F(1, 33) = 3.48$; $P = .071$. The Group × Gender interaction was clearly non-significant, $F(2, 33) = .49$; $P > .05$. The VE group differed significantly from both washing line and caption groups (P's $< .05$ and $.003$ respectively) but the latter groups did not differ, $P = .19$. Removed2 scores showed the same pattern; groups differed, $F(2, 33) = 4.64$; $P < .02$, and a gender difference (male superiority) approached significance, $F(1, 33) = 3.2$; $P = .082$. The VE group was better than either washing line or caption groups, P's $< .05$ and $.005$ respectively, but the latter groups failed to differ, $P = .34$.

The SPE (Fig. 3) was analysed using non-parametric statistics, the Kruskal–Wallace test being used to conduct a one-way independent groups analysis on each successive serial block, group differences being examined for significance using the Mann–Whitney U-test. Group differences were observed in the middle block only (positions 4–6),
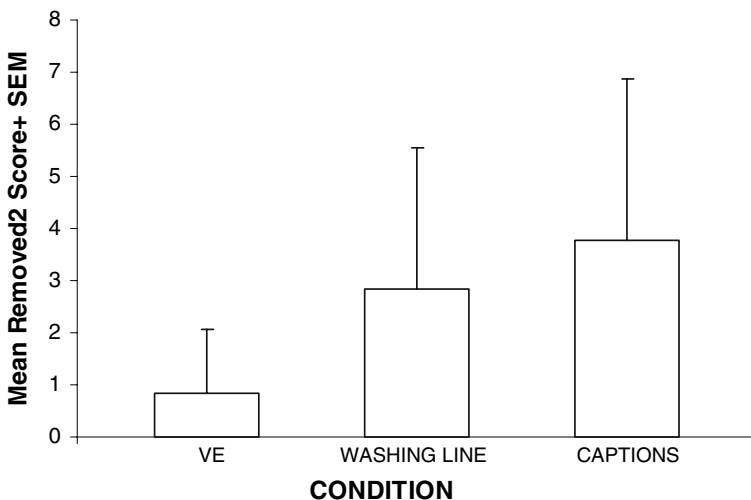

Fig. 2 Mean Removed2 error scores in the VE, washing line and captions groups in Experiment 1
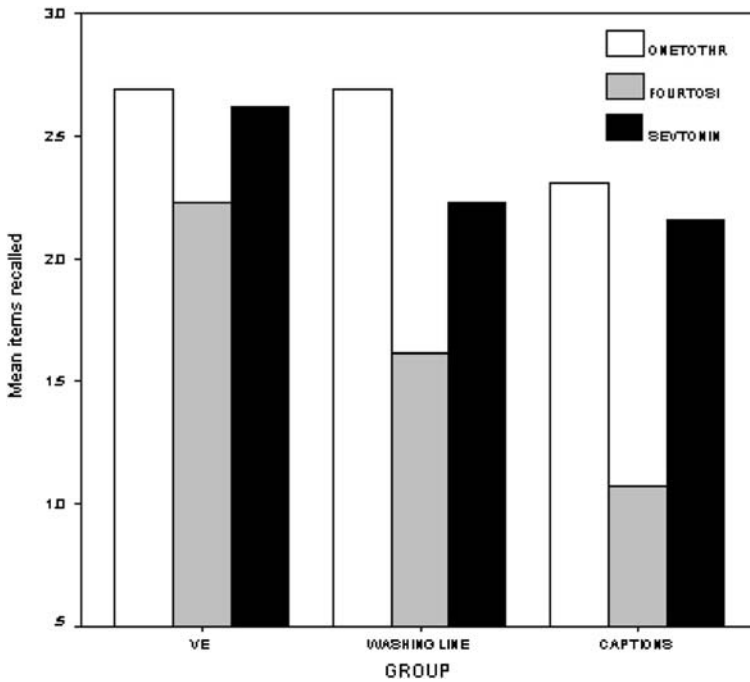
Fig. 3 Serial position effect (blocks of positions 1–3, 4–6 and 7–9) in the three groups in Experiment 1

$\chi^2(2) = 5.91$; $P = .05$, due to higher scores in the VE group compared with the caption group, U(13,13) = .42; $P = .021$, though other group comparisons failed to reach significance.

Discussion

The first study has shown that students experiencing an imaginary historical sequence of events were better able to place those events in order after a virtual presentation than after viewing a 2-D sequence or a series of verbal captions, in particular showing better recall of information in middle list positions. Groups showed significant differences on three out of four measures, and almost reached significance on the number of questions answered correctly. The results support the general assumption that computer technology can improve upon other media in imparting sequential information (Cooper 2000; O'Hara and O'Hara 2001; Wood and Holden 1995; Watson 1993; Yaxley 2004); this may be due to its motivating characteristics (Smart 1996; Underwood and Underwood 1990), although we cannot rule out alternative explanations (greater vividness of presentation, for example). Near significant gender differences, favouring males, were apparent on questions answered correctly, and both Removed scores. Note that all participants were equally unfamiliar with the materials used (since they were fictitious), and that all participants were familiar and relaxed in computer use. The VE presentation was significantly more effective than the washing line, and, in spite of the popularity of 2-D time lines (O'Hara and O'Hara 2001; Smart 1996), the washing line presentation here was notably not significantly better than the use of printed captions alone. The presentation of historical material in these forms

does not appear to engage memory successfully. Although used here with undergraduate students, the control conditions corresponded to modes of presentation that would be used conventionally in history teaching in the classroom (O'Hara and O'Hara 2001; Smart 1996).

Following these encouraging results vis-à-vis the VE condition, its application was explored using standard classroom NC history materials, using a school age group which was trained and tested in their regular history classes. A computer-based 2-D condition was introduced, as a strong control for the non-specific motivation associated with computer use by the VE group.

Experiment 2

Methods

Participants

The participants were 63 children in year 9, 30 male and 33 female, aged between 11 and 14 years, studying history as a NC subject in secondary schools in north London. Two classes participated from one of the schools and one class from the second. Parental consent for participation was obtained. All had normal or correted-to-normal vision.

Materials

A 9-item VE fly-through was constructed as a chronological sequence and presented as in Experiment 1, but incorporating significant events in Feudal England (see Appendix 1b), with captions but omitting dates (since dates could not be accurately ascribed to some events, e.g., the construction of different styles of castle, and three successive events occurred within the year 1066). The events were selected by school history staff to represent materials with which they judged children to experience memorial difficulty. The same nine images with captions were used to create both a PowerPoint 2-D computer graphic presentation [using PowerPoint 2000 software] and a series of graphics on A4 sheets of paper in portrait orientation. For the PowerPoint presentation, as in the VE, a final screen asked if they wished to return to the beginning of the display. Pressing ''Y[es]'' restarted the procedure beginning with the first item in the historical sequence. The specifications of the computers were as for Experiment 1.

Procedure

The era of history targeted in this study began just prior to the Norman conquest of 1066, and encompassing the Norman period, which was particularly characterised by the construction of wooden and later stone castles in the period up to 1300. This was taught as a history NC topic; all participants had been taught about using standard illustrated texts, for 3–5 weeks prior to the start of the experiment. They were trained and tested in this experiment during regular history teaching periods, in groups of 5 in a computer room close to their usual class room. The experimenter spent several days working with the groups and explaining the procedure to be adopted, before starting training.

The study took place in 2 phases, a first phase with a group in school 1 and then a second phase involving new groups of children in both participating schools. Groups were allocated by teaching staff, so that each included a wide range of abilities.

Participants in the VE condition ($N = 26$; 17M, 9F) explored in small groups in phase 1 (4 or 5), each taking a turn in operating the directional key, and individually in phase 2. All participants passed through the VE 5 times. Note that the VE used in the second test phase was somewhat more elaborate than that used in the first, incorporating a number of animated stimuli which were relevant to the nine test stimuli and positioned adjacent to them plus auditory cues (e.g., battle sounds) for some stimuli, in phase 2. However, since the results of phases 1 and 2 were comparable, all data have been combined for statistical analysis. Run-time to pass through the VE was 67 s.; return to the start took 15 s.

For the PowerPoint condition ($N = 13$; 6M, 7F) (only introduced in phase 2) the participant pressed a downward directional keyboard key to move from one image to another. At the end of the nine images, an additional screen offered the opportunity to return to the beginning (as for the VE condition). The paper control condition ($N = 24$; 7M, 17F) involved the children scrutinising the nine images in sequence on successive sheets of paper (see above). In all cases children went through all the materials five times. They were encouraged by the experimenter, who used verbal narratives (emphasising terms such as ''before'', ''after'' and ''then...'') to place the items in context, for all conditions.

Participants were tested individually after a delay of 48 h. Testing began with 9 images as A4 portrait sheets, which they had to place in correct order along a desk. Placement time was unlimited. Participants varied in the time taken, but usually this was between 4 and 8 min. After this, they were presented with the nine questions (''Did X come before Y?''); usually this took approximately 7–10 min.

To test for the possibility that VE learning might have greater durability, one of the test groups ($N = 26$, $N = 13$ in a VE group, $N = 13$ in a paper control group) was retested in full after a vacation, 2 months after the original training.
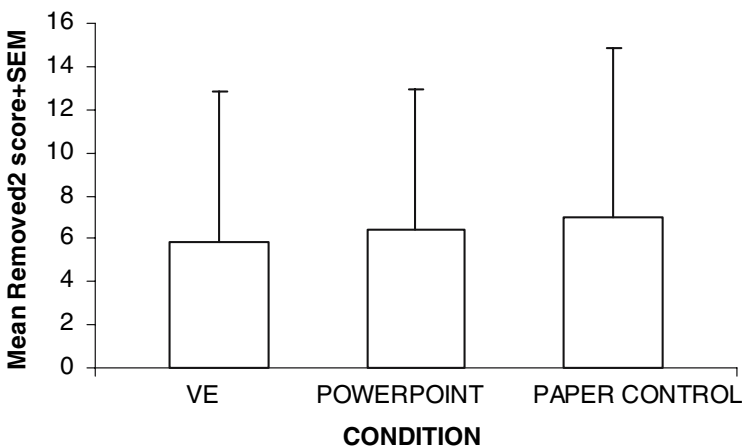

Fig. 4 Mean Removed2 scores in the VE, PowerPoint and paper groups in Experiment 2

Results

Data were analysed as in Experiment 1. In terms of picture ordering, there was no significant effect of Condition, $F(2, 57) = 1.12$; $P = .33$. No significant difference emerged between males and females, $F(1, 57) = .89$; $P = .35$, and there was no Gender × Condition interaction, $F(2, 57) = .19$; $P = .82$. The same pattern emerged for the number of questions answered correctly, and for both Removed1 and Removed2 scores (for Removed1 scores, see Fig. 4). There were no significant group differences when primacy, mid-list and recency position blocks were analysed, $\chi^2(2) = 1.03, 1.18$ and 1.53 respectively; P's > .05.

Results were related to staff ratings of students' ability in history, but there was no overall evidence of a relationship between these measures. In phase 1 there appeared to be a relationship such that the children rated poorer at history placed more items correctly after the VE than in the paper control condition ($P = .052$). However, this effect was not replicated in phase 2 in either sample.

The subgroups tested after a 2 month interval did not demonstrate any significant differences, so the VE group was no better able to bridge the time gap than those in the paper control group. However, a high correlation was obtained between the picture ordering scores at the first test and after the delay (r[df = 24] = .7; $P < .001$).

Discussion

The results of Experiment 2 were disappointing, insofar as the VE presentation did not confer any memorial advantage compared with the presentation of the same materials graphically on paper sheets (as would often be done in the course of history teaching) or as a PowerPoint presentation of successive 2-D graphical images. At least, we must conclude that there was no benefit of using a VE to supplement standard teaching classes, in which students had previously encountered the materials. In addition, there was no benefit of using a VE in terms of the longevity of memory.

In this study, the pictures used in training were also those used in testing, which may have given the paper control group an advantage (though verbal captions alone would have been invalid as a control in this study, since classroom teaching of the relevant materials involved the use of images with captions and dates). In addition, the materials used may not have been sufficiently exciting to construct a really engaging VE (Norman kings, a battle, invasion and castles), so that secondary age pupils were not sufficiently motivated. An explanation in terms of the unreliability of the test protocol is unlikely in view of the high reliability of picture ordering performance across two test sessions in a sample of participants.

A final study employed a younger age group (in a primary school) who were encountering NC history materials that they had not yet been taught. We hypothesised that the use of VEs compared with 2-D media ought to result in better retention, due to the motivation created by computer and VE use, unless overexcitement diminishes recall. Use of paper graphics was expected to be the least successful medium. Necessarily, coarser eras of history were used from which exciting images could be constructed.

Experiment 3

Method

Participants

Seventy-two primary school children (39 male, 33 female) participated in the present study, recruited via their school on a voluntary basis, 35 children in year 3 (19M, 16F, 7–8 years old) and 37 children in year 4 (20M, 17F, 8–9 years old). All had extensive previous experience of using a computer keyboard. They had normal or corrected-to-normal vision. Consent for participation was obtained from parents.

Materials

The stimuli were nine images, selected by school history staff from classroom NC history materials, to represent historical eras with which they judged children to experience memorial difficulty. Each was dated using a particular date, for simplicity, rather than a date spread, and was labelled with a brief description of the historical event or period as a caption passing from 3000 BC to 1939 AD (Appendix 1c). Ancient Egypt was charac- terised by a photograph of pyramids, plus caption and date, while evacuation during WWII was represented by a photograph of children boarding a train with suitcases, plus caption and date, also an added 3-D feature, a noisy hurricane aircraft that flew overhead.

The VE fly-through was presented and controlled as Experiments 1 and 2. It was run on an RM Nimbus PC with Windows 98. In the PowerPoint condition the materials were presented as a slide show running on the same computer as above; presentation and key operation were as in Experiment 2. In the paper condition the same image sequence was presented on sheets of A4 paper, two on each of pages 1–4, and one on page 5. These pages were securely fastened by a staple in the top left hand corner.

Procedure

The participants were divided into three independent groups by their class teachers. The teachers were asked to divide the children up evenly with a mixture of boys and girls and a spread of abilities across the conditions. Training occurred in groups in a computer room close to the children's classroom.

For the VE condition participants ($N = 24$; 13M, 11F) were instructed on how to travel through the VE by depressing the number 8 key on the right hand side of the keyboard. They were instructed to study the screens as they travelled past them, as though they were walking through the environment, and to try to remember as much information as possible. They were prompted at the end of each fly-through to return to the beginning again. Run- time through the VE was 62 s.; returning to the start took 20 sec.

The PowerPoint condition ($N = 23$; 12M, 11F) was run as in Experiment 2. For the paper condition ($N = 25$; 14M, 11F), thirteen sets of stapled sheets were printed and distributed to the participants who were asked to study the sheets in order. In all conditions participants were asked to study the materials carefully, remembering as much information as possible about the order of events, and returning to the start when the series had been completed.

In all three conditions, participants spent 20 min studying the materials, free to interact with others. The experimenter and participants read aloud the dates and descriptions in the first run through the materials, to clarify any unknown words or dates. The chronological order of events and dates was specifically brought to the participants' attention (e.g., that the ancient Egyptians pyramids were the first image to be encountered, because they were the farthest back in history). The teacher and experimenter circulated among the group, encouraging them and answering questions. Participants were unaware that they would later be tested. All other procedural details were as for Experiment 2.

Testing took place 2 days after the training session. Participants were tested individually for 5–10 min depending on how long they took to answer the questions. They were asked to place the 9 images (on A5 laminated card) in their correct order on a flat desk surface. Nine questions were then presented on an A4 size sheet of paper, relating to the chronological order of the pictures (''Did X come before Y?'' or ''What event came between events Y and Z?'' scored as correct or incorrect).

Results

Dependent variables and analyses were as in Experiments 1 and 2.

Teachers rated pupils' academic ability on a 1–3 scale, based on membership of ability ''sets'' within the class (1 = high, 3 = low). The allocation of participants to paper, PowerPoint and VE conditions was random since there was no significant intergroup difference in ability ratings ($\chi^2[4] = 2.77$; $P = .60$). There was no gender difference in teacher-rated ability (U(39,33) = 607; $P = .65$).

For pictures placed correctly in sequence, groups failed to differ, $F(2,66) = 1.38$; $P > .05$; a gender difference favouring boys (mean: 6.72) over girls (mean: 5.58) almost reached significance, $F(1, 66) = 3.85$; $P = .054$, though the Group × Gender interaction was not significant, $F(2, 66) = .094$; $P > .05$.

When data on the number of questions answered correctly were entered into a two way, Group × Gender ANOVA, groups differed significantly, $F(2, 66) = 3.86$; $P < .05$, since the paper group was significantly better than either the VE or the PowerPoint groups (P's = .021 and .019 respectively) though the two latter groups failed to differ, $P = .96$. There was no significant Group × Gender effect, $F(2, 66) = 1.15$; $P > .05$.

Perhaps unsurprisingly, teachers' ability ratings correlated significantly with the number of questions answered correctly (Spearman's ρ [$N = 72$] = .22; P[1-tailed] = .03).

Differences among the presentation groups' Removed1, and Removed2, scores were not significant, $F(2, 66) = 1.8$ and 1.4 respectively; P's > .05. Again teachers' ability ratings were significantly correlated with the participant's performance on Removed1 and Removed2 ($\rho[N = 72]$ = .19 and .20 respectively; both P's[1-tailed] = .05).

Serial order effects were tested as in Experiments 1 and 2. Despite a clear trend in the data (Fig. 5), the Kruskal–Wallis test showed no significant group differences when comparisons were made within individual blocks, though when data from the first 2 blocks were combined, placement accuracy in these list positions (1–6) was significantly better in the paper group than in the two computer groups combined, U(25,47) = 423; $P = .046$. Boys (mean = 2.67) showed a superiority over girls (mean = 2.18) in primacy positions 1–3, U(39,33) = 406; $P = .01$, though no such superiority was evident for recency positions 7–9, U = 581; $P = .43$.

To determine the possibility that lower ability children benefited most from experiencing a VE presentation, data were reanalysed using only the scores from
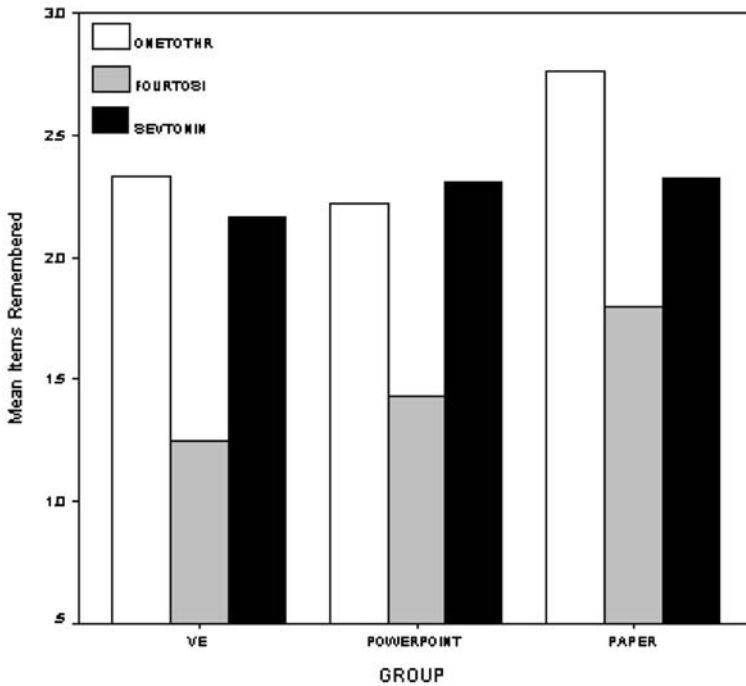
Fig. 5 Serial position effect in the three groups in Experiment 3

participants with ability levels 2 or 3 (N's = 10, 14 and 13 for the VE, PowerPoint and paper groups respectively). However the elimination of the high-flying children made no difference to the group comparisons described above, on any dependent measure.

Discussion

Participation in a particular experimental group was not a strong predictor of performance, and was surpassed by teachers' ratings of children's general academic ability (based on classroom performance), in terms of the number of questions answered correctly. Indeed, in terms of all four dependent measures used, VE use had a detrimental effect on the retention of chronological sequencing. This seems to be so especially for items in early/intermediate list positions. Our results suggest either that memory in children of this age is especially vulnerable to influences that produce loss of information in the earlier list positions, possibly due to the over-stimulating effect of using an exciting medium, Children in the VE condition may have been prevented from benefiting if they lacked true familiarity with the spatial environment, and thus were still remembering the list of events as a list of semantic items rather than coding them as being located in successive spatial locations.

Interestingly, there was a gender difference in terms of SPE, boys showing a superiority in early serial positions, which, to our knowledge has not been previously reported.

General discussion

While the results of Experiment 1 seemed to support the optimism of Passig and Sharbat (2000) that VE technology could revolutionise teaching and learning, by replacing traditional teaching methods (improving the recall of historical information, and enhancing the chronological sequencing of events), the results of Experiment 2 suggested that in terms of history NC materials, this optimism is not justified. Indeed, in Experiment 3, it was found that primary age children suffer a small disadvantage when using computer technology to learn historical material, since their performance in answering chronological questions correctly was better after seeing successive images on paper than after VE and PowerPoint presentations.

The three experiments reported here differed in terms of protocols, the 2-D control conditions used, materials used to create the displays, and other variables. This was necessary and inevitable, in view of the types of material that each age group encounters in the course of their different levels of knowledge and history teaching. The tailoring of environments according to the types of information being used was in keeping with the heuristic nature of the current study, though it limits the comparability among the three reported studies. Nevertheless, to the extent that inter-experiment comparisons are possible, any VE benefit for older participants may have arisen because they were less distracted by the arousing effect of using 3-D technology; they had had extensive experience of computer and 3-D game use. Younger pupils, however, appeared to the experimenter to have been over-aroused by using a computer medium, and this was perhaps reflected in their poorer performances in the computer conditions. Mild arousal can increase recall of materials in intermediate list positions (see Barraclough and Foreman 1994), while a higher level can reduce information acquisition, which may explain the selectively depressed retention of information from early and intermediate list positions in Experiment 3. Clearly, specific information contained in the presentations may have influenced the results; a hurricane aircraft flew over the final item in the VE display in Experiment 3 which may have been especially exciting, though this does not explain the superiority of participants in the paper group over those in *both* the VE and PowerPoint groups. Clearly, it is necessary to conduct studies in which early and intermediate list information is rendered especially exciting to determine the relationship between item position and interest. Children could begin the fly-through at differing points, to reduce serial order effects.

Except in Experiment 1, where inter-event intervals could be equalised, the fly-throughs in Experiments 2 and 3 did not attempt to represent ''time elapsed'' between events with proportional time intervals in the display. This would have resulted in long waiting times between some events (which could have depressed memory). Children have specific difficulties with both sequencing and the duration of time between different events (Wood and Holden 1995). This should be considered when historical events are selected for inclusion in a virtual sequence, but timelines can also potentially build on childrens' personal historical knowledge of recent events and their spacing (cf. Friedman 1982) to establish the principle before encountering historical materials; this can enhance childrens' ability to use terminology which is relevant to describing time, such as ''before'', ''after'' and eventually to utilise more precise units of time (Smart 1996).

Children may benefit from active involvement in the construction of their own time lines (cf. Pedley et al. 2003), by actively pasting in relevant pictures and textual information that they find in other historical packages and sources. Elements of the timeline can

then be rearranged and readily modified. ICT offers a range of possibilities; timelines have been used in a variety of ways, for example allowing children to view the past histories of both England and Scotland from the Neolithic period using pictures and three dimensional models (see Cooper 2000). Concepts such as AD-BC could potentially be addressed using a fly-through that flags the fact that the year 0 is approaching, then passed. (The historical sequence in Experiment 3 passed from BC to AD dates, though children did not appear confused by this). Furthermore, test protocols can be developed in VEs which are impossible in other media; participants may be asked, in a game format, to guess the content of blank screens as they are approached, feedback being provided via the appearance of the (correct) screen image at a preset virtual distance.

An absence of any VE group advantage in Experiments 2 and 3 may have occurred because participants had not become sufficiently familiar with the VEs during training for the successive stimuli to represent ''places'' in spatial memory. This study attempted to make users in VE conditions feel as though they were part of the environment, as if they were walking along a familiar shopping street, making the user feel ''present'' by incorporating elements which Kullberg (1995, 1996) suggests makes the visual space feel 'real'. Children claimed to be familiar with the spatial environments after passing through the VEs on five occasions. Nevertheless, in studies involving exploration of very large environments, 2 h of exploration has been used in earlier studies (e.g. Foreman et al. 2003, 2005), though a few minutes of exploration can be adequate to impart good spatial information (cf. Waller 2000). Future studies using game formats might engage attention such that longer training intervals are tolerable.

It is also important to know whether the use of familiar real or VEs (e.g., actual or modelled school rooms and corridors; see Foreman et al. 2003) may be an effective way of tagging places to time-space information that the teacher wishes pupils to learn (not only in history but other subjects with an historical context such as sociology, religion, art or music). Further studies are required that train children in the understanding of the spatial layout prior to the introduction of the historical events to-be-remembered.

Clearly, the use of spatial memory does not demand the use of computer technology, though evidence suggests that VEs are suitable to convey such information (Bliss et al. 1997; Foster et al. 1998; Foreman et al. 2003, 2005; Tate et al. 1997). As with IMM techniques (Masterman and Rogers 2002) and 3-D spatio-temporal environments (Kullberg 1995, 1996), it is important to test the use of these media against conventional benchmarks. The present study contributes to a wider debate, concerned with the benefits of adopting new technologies in teaching (whether VE-based or PowerPoint), and questions the assumption that electronic technology is necessarily advantageous compared with conventional approaches (see also Haydn 2002, 2004, 2006). An important conclusion from the present results is that, despite the potentially motivating influence of technology (Smart 1996; Underwood and Underwood 1990; Watson 1993), the use of traditional methods in history teaching such as the use of paper images can be as effective, and sometimes more effective, than technological media.

Appendix 1a, b, c

Content of fly-throughs in Experiments 1–3

(a) Imaginary planet

Aliens land on Planet Gozon 2080
Massive monuments erected to honour moon god 2100
The crowning of the Great King 2120
Devastating eruption of fire mountain 2140
The building of the Grand Palace 2160
The Great Battle 2180
The crowning of Queen Vanessa 2200
The Great Flood of 2220
The Fire of 2240

(b) Feudal England

Edward the confessor is king
Edward sends Harold on a mission
Death of Edward the confessor
Harold the King is crowned
Battle of Hastings
King Harold is killed
William the Conqueror crowned
Wooden Castle building begins
Stone castle building begins

(c) General history

Ancient Egypt-Egyptian Pyramids, 3000 BC
Ancient Greek Mythology-Theseus and the Minotaur, 1500 BC
Celts' round houses in small villages, 600 BC
Roman Centurion, 200 BC
Helmets worn by Anglo Saxons, 449 AD
Viking ship (long boat) used to travel along rivers, 787 AD
Normans—Battle of Hastings, 1066 AD
The Great Fire of London, 1666 AD
Evacuees leave London in World War 2, 1939 AD

References

Atkinson, R. C., & Shiffrin, R. M. (1971). The control of short-term memory. *Scientific American, 225*, 82–90.

Barcelo, J. A., Forte, M., & Sanders. D. H. (2000). *Virtual reality in archaeology*. Oxford, UK: Archaeopress.

Barraclough, S., & Foreman, N. (1994). Factors influencing recall of supraspan word lists: Caffeine dose and introversion. *Pharmacopsychoecologia, 7*, 229–236.

Bliss, J. P., Tidwell, P. D., & Guest, M. A. (1997). The effectiveness of virtual reality for administering spatial navigational training to firefighters. *Presence, Teleoperators and Virtual Environments, 6*, 73–86.

Bourdillon, H. (1994). *Teaching history*. London: Routledge.

Broadbent, D. (1981). *Decision and stress*. Oxford: Oxford University Press.

Cooper, H. (2000). *The teaching of history in primary schools: Implementing the revised national curriculum* (3rd ed.). London: David Fulton Publishers.

Friedman, W. J. (1982). Conventional time concepts and children's structuring of time. In W. J. Friedman (Ed.), *The developmental psychology of time* (pp. 171–208). New York: Academic Press.

Foreman, N. P., & Gillett, R. (Eds.). (1997). *Handbook of spatial research paradigms and methodologies, Volume 1: Spatial cognition in child and adult*. Hove: Psychology Press.

Foreman, N. P., & Gillett, R. (Eds.). (1998). *Handbook of spatial research paradigms and methodologies, Volume 2: Clinical and comparative issues*. Hove: Psychology Press.

Foreman, N., Stanton, D., Wilson, P., & Duffy, H. (2003). Spatial knowledge of a real school environment acquired from virtual or physical models by able-bodied children and children with physical disabilities. *Journal of Experimental Psychology: Applied, 9*, 67–74.

Foreman, N., Stanton, D., Wilson, P., Duffy, H. & Parnell, R. (2005). Transfer of spatial knowledge to a 2-level shopping mall in older people, following virtual exploration. *Environment and Behavior, 37*, 275-292.

Foster, G. T., Wenn, D. E. N., & Harwin, W. S. (1998). Generating virtual environments to allow increased access to the built environment. *The International Journal of Virtual Reality, 3*, 12–20.

Glanzer, M., & Cunitz, A. R. (1966). Two storage mechanisms in free recall. *Journal of Verbal Learning and Verbal Behavior, 5*, 351–360.

Haydn, T. (1999). Teaching children about time; back to basics? In R. Phillips, & G. Easdown (Eds.), *History education: Subject knowledge, pedagogy and practice* (pp. 117–130). Lancaster: Standing Conference of History Teacher Educators in the United Kingdom.

Haydn, T. (2002). The book versus the screen: Educational media in the digital age. *Paedagogica Historica: International Journal of the History of Education, 38*(1), 387–401.

Haydn, T. (2004). What do they do with the information? Computers in the history classroom: Some lessons from the United Kingdom. International Experience of ICT usage in education. In *Proceedings of the 12th UNESCO/IITE International Conference on Information Technologies in Education, 2002*. (pp. 59–78). Moscow: UNESCO/IITE.

Haydn, T. (2006). Multimedia, interactivity and learning: Some lessons from the UK. In *Proceedings of the 4th International Conference on Multimedia and ICT in Education*. Seville, Spain, November 22–25.

Hodkinson, A. (1995). Historical time and national curriculum. *Teaching History, 79*, 18–20.

Hoodless, P. (1996). *Time and timelines in the primary school*. London: The Historical Association.

Jacobs, W. J., Laurance, H. E., & Thomas, K. G. F. (1997). Place learning in virtual space I: Acquisition, overshadowing and transfer. *Learning and Motivation, 28*, 521–541.

Kullberg, R. L. (1995). Dynamic timelines: Visualizing historical information in three dimensions. MSc. Thesis, MIT Media Laboratory. Boston, MA: MIT.

Kullberg, R. L. (1996). Dynamic timelines: Visualising the history of photography. *Videos, CHI 96*, 386–387.

Madeley, H. (1921). Time charts. Pamphlet No. 50, Historical Association.

Masterman, E., & Rogers, Y. (2002). A framework for designing interactive multimedia to scaffold young children's understanding of historical chronology. *Instructional Science, 30*, 221–241.

Miller, G. A. (1956). The magical number seven plus or minus two: Some limits on our capacity for processing information. *Psychological Review, 63*, 81–97.

O'Hara, L., & O'Hara, M. (2001). *Teaching history 3–11: The essential guide*. London: Continuum.

O'Keefe, J., & Nadel, L. (1978). *The Hippocampus as a cognitive map*. Oxford: Clarendon Press.

O'Neil, M. J. (1992). Effects of familiarity and plan complexity on way-finding in simulated buildings. *Journal of Environmental Psychology, 12*, 319–327.

Partington, G. (1980). *The idea of an historical education*. Windsor: NFER Publishing Company.

Passig, D., & Sharbat, A. (2000). The why and how of VR in schools: A preferred future pedagogic mission by a group of worldwide experts in VR and education. International Journal of Virtual Reality, http://www.passig.com/pic/VRinSchool.htm, pp. 1–7.

Pedley, J., Camfield, L., & Foreman, N. (2003). Navigating memories. In B. Ahrends, & D. Thackara (Eds.), *Experiment: Conversations in art and science* (pp. 173–235). London: Wellcome Trust.

Qualifications and Curriculum Authority (QCA). (1999). *The national curriculum: Handbook for primary teachers in England*. London: DfEE/QCA.

Richardson, A. E., Montello, D. R., & Hegarty, M. (1999). Spatial knowledge acquisition from maps and from navigation in real and virtual environments. *Memory and Cognition, 27*, 741–751.

Ruddle, R. A., Payne, S. J., & Jones, D. M. (1997). Navigating buildings in ''desk-top'' virtual environments: Experimental investigations using extended navigational experience. *Journal of Experimental Psychology: Applied, 3*, 143–159.

Scaife, M., & Rogers, Y. (1996). External cognition: How do graphical representations work? *International Journal of Human-Computer Studies, 45*, 185–213.

Smart, L. (1996). *Using I.T. in primary school history*. London: Cassell.

Stanton, D., Wilson, P., & Foreman, N. (1996). Using virtual reality to aid spatial awareness in disabled children. In P. Sharkey (Ed.), *Proceedings of the First International Conference on Disability, Virtual Reality and Associated Technologies* (pp. 93–101). University of Reading.

Tate, D. L., Silbert, L., & King, T. (1997). Using virtual environments to train firefighters. *IEEE Computer Graphics and Applications, 17*, 23–29.

Tolman, E. C. (1948). Cognitive maps in rats and men. *Psychological Review, 55*, 189–208.

Turner, A. (1998). ''It would have been bad''. The development of historical imagination and empathy in a group of secondary age pupils with severe learning difficulties. *British Journal of Special Education, 25*, 164–166.

Underwood, J., & Underwood, G. (1990). *Computers and learning*. Oxford: Blackwell.

Waller, D. I. (2000). Individual differences in spatial learning from computer generated environments. *Journal of Experimental Psychology: Applied, 6*, 307–321.

Watson, D. (1993). *The impact report: An evaluation of the impact of educational technology on children's achievements*. London: Kings College.

Wilson, P. N. (1997). Use of virtual reality computing in spatial learning research. In N. Foreman & R. Gillett (Eds.), *Handbook of spatial research paradigms and methodologies: Volume 1, Spatial cognition in the child and adult* (pp. 181–206). Hove: Psychology Press.

Wood, D., Bruner, J. S., & Ross, G. (1976). The role of tutoring in problem solving. *Journal of Child Psychology and Psychiatry, 17*, 89–100.

Wood, L., & Holden, C. (1995). *Teaching early years history*. Cambridge: Chris Kington Publishing.

Yaxley, J. (2004). Experiencing history. *Teaching History, 42*, 24–25.