

This item was submitted to Loughborough's Institutional Repository by the author and is made available under the following Creative Commons Licence conditions.



creative commons
COMMONS DEED

Attribution-NonCommercial-NoDerivs 2.5

You are free:

- to copy, distribute, display, and perform the work

Under the following conditions:

 **Attribution.** You must attribute the work in the manner specified by the author or licensor.

 **Noncommercial.** You may not use this work for commercial purposes.

 **No Derivative Works.** You may not alter, transform, or build upon this work.

- For any reuse or distribution, you must make clear to others the license terms of this work.
- Any of these conditions can be waived if you get permission from the copyright holder.

Your fair use and other rights are in no way affected by the above.

This is a human-readable summary of the [Legal Code \(the full license\)](#).

[Disclaimer](#) 

For the full text of this licence, please go to:
<http://creativecommons.org/licenses/by-nc-nd/2.5/>

The Effect of a Computer-Based Cartooning Tool On Children's Cartoons and Written Stories

M. Madden*, P.W.H. Chung, C.W. Dawson.

*Department of Computer Science, Loughborough University, Loughborough,
LE11 3TU, UK*

Keywords: teaching/learning strategies, media in education, improving classroom teaching, evaluation of CAL systems, human-computer interface

Abstract

This paper reports a study assessing a new computer tool for cartoon storytelling, created by the authors for a target audience in the upper half of the English and Welsh Key Stage 2 (Years 5 and 6, covering ages 9 to 11 years). The tool attempts to provide users with more opportunities for expressive visualisation than previous educational software; its design was motivated by earlier work connecting “moving image literacy” with print literacy, and it was used here in storywriting preparation work: users first visualised a known story, then wrote their versions of it based on the cartoons they had made. The stories produced are compared with stories written using two other preparation activities, one a pencil-and-paper cartooning exercise and the other a teacher's normal planning session, which also resulted in a retelling of a known story.

The study finds that no one preparation process had a noticeably different effect on the final written stories; however, the cartoons produced with the software are found to be quite different to their paper counterparts, showing a greater variety of character action, pose and interaction, slightly more variety of camera shot distance, and more pictures. Children's and teachers' reactions to the software tool are also discussed.

1 Introduction

In recent years, the concept of “moving image literacy” as a skill which may be learned alongside, and perhaps reinforce, conventional notions of English literacy has been described a number of times (see, for example, Parker, 1999; Oldham, 1999; Higgins, 2002). Parker (1999) and others, by building on Robinson (1997), assert that children can benefit from considering the ways in which the same narrative concepts, such as characterisation, setting, pace and so on, are expressed differently in the moving image and the written word, and that children's acquisition of literacy concepts can be scaffolded by translating the techniques of one medium into another; in some ways this echoes work in other areas, for example the use of visual ‘advance organisers’ (Chun and Plass, 1996) in second language acquisition. Moving image media are cited particularly for their familiarity amongst even pre-school children, and for the possibility that learning activities which incorporate them may help to connect school life with the wider world (Marsh and Thompson, 2001).

*Corresponding author. *Email address:* mdm103@googlemail.com, *Fax:* +44 115 846 1153

Interest in this use of the moving image is reflected in the UK Government's aspirations for English teaching in primary schools; the 2004-2005 annual report of Her Majesty's Chief Inspector of Schools (Office for Standards in Education (Ofsted), 2005) articulates a specific desire for more imaginative work using moving image "texts" in the English curriculum. However, as Burn and Leach (2004) note, it is important to recognise the danger of over-using language analogies when referring to "texts", "literacies" and so on; really, the moving image uses a more general system of signification and is not precisely a language. This brings definitions into focus. Although informally, it is easy to identify certain things - film, television, animation and suchlike - as employing the moving image, is it possible to make a more fundamental statement as to what a "moving image text" really entails?

Burn and Parker (2003a; 2003b; 2001) identify two key characteristics in any such text: *filming* and *editing*. Filming requires that a visual scene has been framed, and its elements composed, so that it is seen in actual or notional camera shots. Editing refers to making meaning by placing these shots into a particular order so that their juxtapositions and transitions may create rhythms or set, fulfil or confound expectations. They state that these characteristics together define the *mode* of the moving image, which they name the kineikonic mode. According to Burn and Leach, this "operates as a combinatorial mode which assembles and integrates other modes (speech, image, gesture, music) through its own 'grammar' of filming and editing".

There is no prescription here as to which other modes must be integrated by this 'grammar', nor how filming and editing are brought about; there may be a real, time-based, physical camera as in film, an interaction metaphor presented by software, or even an implicit point-of-view shown in a comic strip, and the 'shots' may be sequenced temporally or spatially. This allows the definition of a moving image text to be interpreted quite generally. Film and television are perhaps the most obvious manifestations but the definition can extend beyond them to such things as comic strips (McCloud, 1994, draws strong parallels between comics and film), computer games (Burn and Parker, 2003a) and storyboards; perhaps counterintuitively, the mode of the moving image meaningfully extends to artefacts which are not themselves "moving".

Using this generality, Madden (2007) makes a survey of current educational ICT products which allow kineikonic storytelling to be performed. This concludes that, so far as character depiction is concerned, current software tends to sit on either side of a "content-centred/build-centred" divide. The former emphasises support for depiction through presupplied but relatively inflexible resource banks like clipart, and the latter emphasises a greater range of depiction at the expense of support, for example by expecting the user to create their own artwork from scratch. The inference is drawn that there is scope to bridge this divide through a new approach to characterisation termed the "enhanced-frontal-orientation" method. A key motivating concept behind this is enabling characters to be depicted in a wide range of poses and expressions without having to generate each one from scratch, and this might be compared with the way in which word processors have been said to relieve the 'drudgery' associated with revision in writing (Cunningham, 2000). Thus a tool is proposed which integrates this approach with simple camera functionality, necessary if kineikonic notions of filming are to be allowed, and a form of process help adapted from the HARRY system described by Holdich and Chung (2003).

The full proposal and its early tests are described in Madden (2007). The current paper will provide evidence that at least some parts of that proposal can allow greater visual expression than previous tools or techniques; it aims also to contribute to the ongoing debate concerning the relationship between ICT, moving image literacy and English teaching (Burn and Leach, 2004; bfi Primary Education Working Group, 2003). It describes a subsequent study designed to assess a later implementation of the software proposal, and to answer three questions concerning its impact in a classroom setting: does the software have a different effect on written stories compared to other processes, does it lead to different kinds of visuals than a pencil-and-paper process, and how much engagement do the users show with the software and the tasks?

2 Software Design

The software used in this study aims to allow users to create a sequence of still images (“shots”) depicting the events in a particular story or story theme; this is accomplished by providing a set of relevant background images, object artwork and character artwork which can be combined to build up each picture. In contrast to normal, static clipart, the character artwork may be ‘posed’ via one interaction mode which allows limbs in a skeleton to be dragged with the mouse, and ‘warped’ via another interaction mode which allows simple manipulation of vector control points. Originally, it was envisaged that users would be able to make their own versions of such characters. However, the software framework written for creating characters was judged unsuitable for novice users, and time constraints prevented it from being rewritten in a friendlier form, resulting in the reliance on pre-supplied characters which were created by the primary investigator.

Figure 1(a) shows the main interface of the software, from which all work is done. The largest area is given over to the current shot being edited; controls for changing the camera’s point-of-view and for adding and editing characters and clipart are provided at its right-hand-side; in addition to being posed and warped, here, characters can also be repainted in different colours. This shot’s place in an explicit left-to-right sequence is shown in the sequence bar, which contains controls for adding, removing and reordering shots, and for selecting different shots to edit. The controls at the left of the sequence bar allow work to be saved and loaded, undo/redo to be performed, and a full-screen slide show of the sequence to be viewed.

Process support is supplied via what are termed “visual storytelling prompts”, accessible via the labelled button at the bottom right of the screen. When this is pressed, a secondary window is displayed which presents a summary of the key stages in the story theme being used; this may be seen in Figure 1(b). In addition, for each story stage, three subsidiary help prompts may also be displayed (Figure 1(c)); “ideas help” gives suggestions for ideas to include in the sequence, “character help” highlights important characterisation aspects to show for each part of the story, and “camera help” suggests particular camera shots or combinations which may be effective. Table 1 shows an example set of visual storytelling prompts tailored for work with Aesop’s fable of the Fox and the Crow.

No typing interface for making captions is provided; the software aims to

emphasise purely visual aspects of communication, to see how these influence and elicit subsequent written work. This decision owes much to practical constraints concerning how the storywriting was to be accomplished in this study; both schools participating here had a single large ICT suite which was shared between all classes, and teacher consultation reinforced the impression that using the computers for storywriting would take too much extra time out of the suites' already crowded schedule.

3 Process and experiment design

The study described here consists of two sub-studies, the first carried out over September and October 2005 in a Nottinghamshire primary school, and the second in November and December 2005, in a Leicestershire primary school. Both schools have above average standards in English teaching according to the most recent Ofsted reports. The sub-studies are depicted in Figure 2; as can be seen, each contains an experimental and a control group. The overall process followed by each group was intended to follow a three-step pattern, beginning with the collection of a first story, produced under entirely normal conditions before the investigator or the study were introduced to the participants. There then follows an intervention designed to encourage participants to prepare to retell a story; after this intervention concludes, the written version of this story is collected and analysed to see if the interventions produced different results. The number of stories indicated may seem smaller than expected given that two entire classes took part; however, the numbers indicate only those participants who produced both first and second stories.

The earlier sub-study was designed to fit into class work on India and the Hindu sacred story, The Ramayana; the subject selected for the intervention and second story involved a section of this. The later sub-study grew out of the class's 'writing worlds' work and resulted in an intervention and story version of the spooky poem *The Visitor*, by Ian Serraillier.

3.1 Ramayana Sub-Study

Participants here came from one of the school's two mixed Year 5/6 classes, covering ages 9-11 years; they produced a first story in completely normal conditions, consisting of the opening part of *The Ramayana*. Beginning the following week, participants divided into experimental and control groups and focused on a later section. The control intervention activity was chosen to be a modified version of a process the teacher had carried out with a previous year, in which students made a captioned comic strip version of a Greek fable. Here, it was agreed that the control group would draw a comic strip without captions, telling the story in mime so as to make the activity as close as reasonably possible to the software one. In the second week, both groups were given a short camera language briefing which consisted of analysing a sequence of eight still clips from a television cartoon, looking at how different shot distances were employed and how the sequence of clips told a story. After two afternoons of their respective intervention work, spread over two weeks and totaling three hours per group, the whole class came together to write the second story in their classroom. Experimental participants were given printouts of their computer cartoons, and

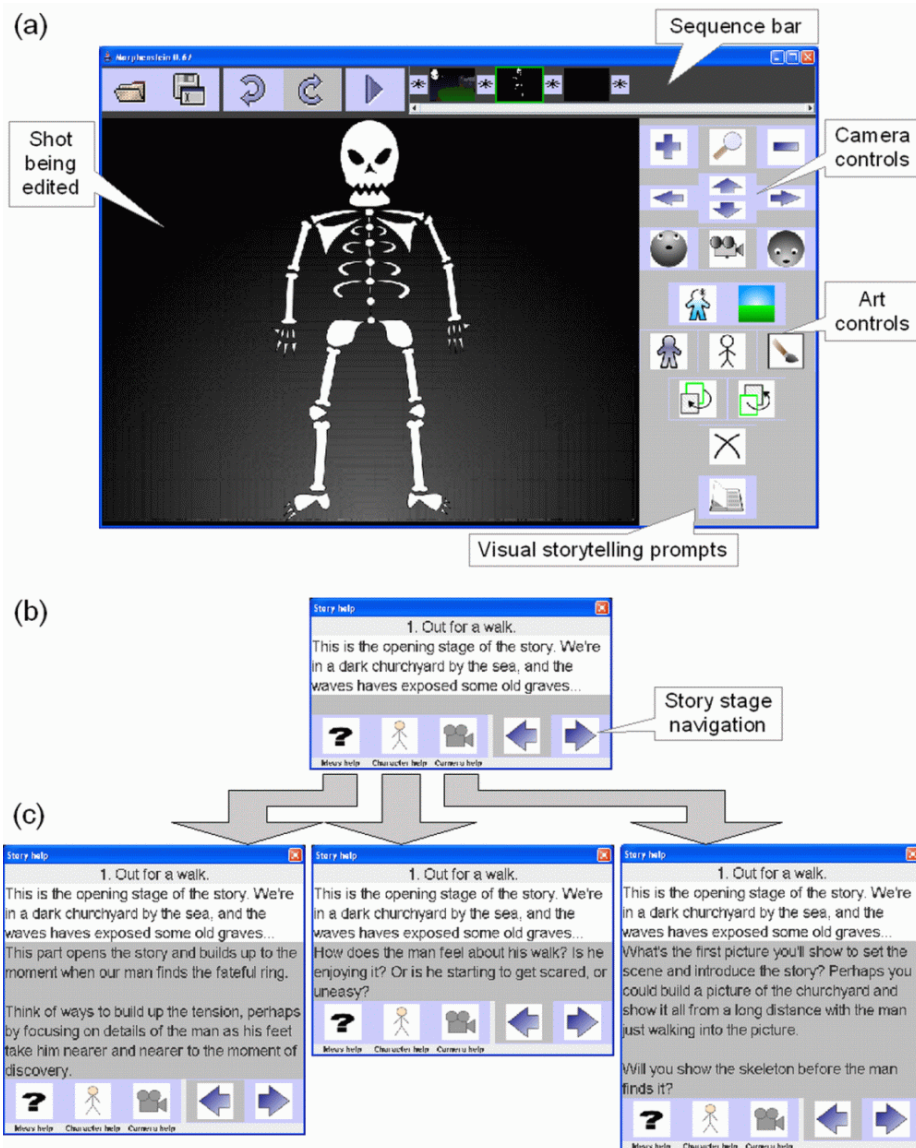


Figure 1: The software used in the study.

Table 1

Visual Storytelling Prompts for Aesop's Fable of the Fox and the Crow

Story stage and guiding prompt	Ideas help	Character help	Camera help
<p>Crow with cheese</p> <p>The first character to be introduced is the Crow. She's holding a piece of cheese in her beak, and she's looking forward to eating it.</p>	<p>Where is the action taking place? Can you set the scene before the Crow appears?</p>	<p>How does the Crow feel about having the cheese? Does she look forward to eating it? How can you show this in her expression and what she does? Is she excited, hopping about? For example, does she rub her hands (or wings) together?</p>	<p>Your first picture could show a forest, or the tree where the Crow is sitting. Once you've shown that, you could use another picture to focus in on the Crow and her cheese. Or you could use more pictures to focus in on the cheese and the Crow separately. Perhaps your last picture could show the Fox appearing.</p>
<p>Crow is conned</p> <p>This is the part where the Fox pretends to admire the Crow.</p>	<p>What does the Fox pretend that he likes about the Crow? Her tail? Her wings? Her beak?</p>	<p>What expression does the Fox have while he's praising the Crow? Does he gesture grandly, or point at anything? How does the Crow react?</p>	<p>Try to make a picture which shows both the Crow and the Fox. Perhaps we could be looking up at the Crow over the Fox's shoulder, or down at the Fox over the Crow's shoulder. For another picture, perhaps you could focus in on some part of the Crow which the Fox is praising.</p>
<p>The Fox gets the cheese</p> <p>There are two important bits to the story here: the Crow dropping the cheese when she opens her mouth, and the Fox getting the cheese. Think about how you'll show these two events happening. You also need to bring the story to a close after the Fox has succeeded.</p>	<p>Is the Crow sad, angry, or something else? How does the Fox feel about getting the cheese?</p>	<p>How do both characters react to the cheese falling to the ground? What expressions and gestures do you think they'll use?</p>	<p>The moment when the Crow opens her mouth and drops the cheese is an important one in the story. Perhaps you could really focus in on this moment. Maybe you could make a sequence showing each character in turn as they watch the cheese fall. What's the last picture you'll show in the story? The Fox, walking away happily? The Crow? Something else?</p>



Figure 2: Overall structure of the study and its constituent sub-studies.

control participants were given their paper cartoons, which they could refer to whilst writing. The first story was written on 20 September, and the second on 5 October. After the second story had been written, the control group were also given time to use the software.

3.2 Visitor Sub-Study

Here, participants originated from the school's Year 5 class. Their first story involved an island rescue, again produced under completely normal conditions. The teacher then gave the experimental group a half-hour software introduction before beginning the story visualisation work. At the time of this software introduction, the class had not yet been told the story of *The Visitor*, so the introduction served only to let the experimental group learn the basic software ropes and did not give them extra time to do the story task.

The control group participants were given a planning sheet on which they could sketch up to eight key scenes of action, and write useful story words under each frame. This was an entirely normal, unmodified activity which was chosen solely by the teacher. In this case, the control and experimental interventions took place over three consecutive days, and totaled a maximum of three and three-quarter hours; the final three-quarters of an hour on the last day were supplied as an optional extra after a software glitch eliminated some experimental participants' first-day work.

As in the *Ramayana* sub-study, groups came together to do their writing after they had completed their respective preparation activities, working from either their paper plans or printouts of their cartoons. The class wrote the first stories on 25 November, and the second on 5 December. Again, the control group also took a turn to use the software after the second stories were complete.

3.3 Hawthorne and Experimenter Effects

Having introduced the experimental structure, it is worth considering two pitfalls which can easily introduce experimental bias. The first is the 'Hawthorne effect', which has been interpreted as an increase in a group's motivation and productivity arising from that group's knowledge that it is participating in an out-of-the-ordinary process (Mayo, 1960, 1975). It is asserted here that, in both sub-studies, any potential for a Hawthorne effect applies equally to all groups. This is so because, firstly, both classes were aware that they as whole classes were doing a piece of work which was being studied by the investigator, and it was made clear that all participants would have a chance to work with the most out-of-the-ordinary part of the exercise, the software, just at different times. This is backed up by comments made by the *Ramayana* sub-study teacher, who noted that the control group were aware that, while they were drawing their paper cartoons, their compatriots were doing a novel task, but this did not demotivate them as they knew that they would have a chance to do that same task. In addition, the *Visitor* sub-study teacher commented that all participants had gone to significant lengths to make a well-decorated final story, which suggests that no one group felt more 'special' or specially-treated than the other. Thus there should be no Hawthorne effect bias in favour of, for example, either experimental group.

The second concerns the ‘experimenter effect’, whereby the leaders of the intervention can (unconsciously or deliberately!) convey the idea that they are interested only in the work of one group of participants. It is suggested that this is addressed here by the fact that both teachers wanted to get good work and good effort out of all students, and so each group was working to an equally relevant and valued finished story. In addition, it was made clear to both teachers that the purpose of the study was to investigate differences in writing rather than showing that, for example, the software resulted in better work.

3.4 Data Analysis

In addition to the experimental structure, Figure 2 also shows the various comparisons which will be applied to the stories and cartoons; they represent specific instances of two general categories: longitudinal comparisons and cross-sectional comparisons. There are four specific longitudinal comparisons to make here, between each of the four groups’ first stories and second stories, to pick up what might loosely be termed an ‘improvement effect’ or a ‘degradation effect’ in each case, depending on whether performance improves or decreases. In addition, five specific cross-sectional comparisons shall be applied. There will be two first-story comparisons, in an attempt to detect any gross prior differences in aptitude between control and experimental groups, and two second-story comparisons, to see if any intervention produced a significantly different effect to its peer. Finally, a cartoon comparison will be made between the software and paper-based cartoons in the Ramayana sub-study, to see if the two media used for cartooning had different effects on the cartoon product. It is not possible to do a similar comparison in the case of the Visitor sub-study; the paper plans are not available, and, even were this not so, they represent rather different animals to the software cartoons. The latter, although they use no text, are capable of standing as complete representations of the story in themselves and represent something finished; all experimental participants in the Visitor sub-study produced a finished cartoon product. By comparison, the paper plans are precisely plans which look to a next stage in production, that is, the written story.

The analysis is broken into two sections, the first relating the first-story, second-story and improvement comparisons and the second to the cartoon comparison. Additionally, a third section considers some qualitative observations.

4 Data Analysis 1: First-Story, Second-Story, and Improvement Comparisons

This section covers the eight comparisons in Figure 2 which deal with participants’ written stories: the first-story comparisons, the second-story comparisons, and the improvement comparisons. These use indicators provided by two analytical tools.

4.1 Choice of Indicators and Analytical Tools

Two complementary tools, described below, are used to gather quantitative data for each participant’s stories. This numerical data is then used as input for

Student's t-tests; the longitudinal comparisons are applied as standard paired t-tests, and the cross-sectional comparisons as unpaired t-tests with equal or unequal variance as appropriate. All tests are performed at the 95% level of significance.

4.1.1 CHECK TEXT: Statistical Features of Prose

The first set of indicators is a set of statistical features for each story, produced by a tool called CHECK TEXT (Holdich, Holdich & Chung, 2002). This is designed to focus on areas of a piece of writing which are of particular interest to teachers. For a given story, it computes 12 different measures ranging from total number of words used and their variety to the percentages of adverbs, common verbs, words over five letters, and use of commas. The full list, together with the tool's own explanatory text, is reproduced in Table 2. The twelve features it describes boil down to eleven actual indicators, because the average sentence length is used in two different ways. All eleven indicators are used, with the following caveats. First, measuring the total number of words may not be very informative, as Holdich, Holdich & Chung (2002) note. Secondly, the measure of average sentence length may not necessarily indicate the variety of sentence lengths used, since a story with uniform sentence length can score the same as a story with mostly uniform sentences, some long, and some short sentences. Lastly, the reciprocal of the sentence to comma ratio is used so as to avoid potential division by zero.

4.1.2 Teacher Coding: Qualitative Impressions

The second set of indicators used arises from each story's qualitative features, as assessed by a qualified teacher unconnected with either sub-study. She was asked to highlight

- **strong imagery** and good **spatial descriptions**;
- good use of **narrative point-of-view**, that is, instances where the storyteller effectively directs the reader's attention towards a significant story element;
- strong **descriptions of characters** or of **actions characters do**.

The first two of these codes are motivated by the benefits implied, for example, by Parker (1999), particularly a greater visuality of work and a greater visualisation of a story's action and events. The last aims to see if the software's flexible characters feed on to any textual description of story characters. The number of occurrences of each code is counted for each story, and these totals used as input for the appropriate t-tests.

The teacher was given a shuffled and unnamed set of 56 stories, consisting of the first and second stories of seven randomly chosen participants from each of the four groups.

4.2 CHECK TEXT Analysis Findings

The findings for each sub-study are considered separately, and conclusions drawn afterwards.

Table 2

The CHECK TEXT Assessment Categories

Total number of Words	A long story indicates that much detail has been included, whilst a short story lacks sufficient detail.
Variety of Words	A low % indicates frequent repetition.
The connective and	And is the most basic connective used by children. A more mature style is achieved by substituting alternative connectives for and or replacing and with a comma or a full stop. A high % of ands indicates immature sentence constructions.
Common connectives (then, so, but, because)	Use of these basic connectives indicates immature sentence constructions.
Simple sentence starters (sentences starting with The, He, She, His, Her, They, We, I)	Sentences which begin with pronouns indicate a weakness in sentence construction - using the same way to start sentences. A low % indicates a more varied writing style.
Average sentence length	Writing should contain a wide variety of sentence lengths. When writing a story, short simple sentences should be interspersed between longer more complex constructions. Short sentences can add suspense or impact. Long sentences can be used for description.
Word length (% of words longer than 5 letters)	A high % indicates an adventurous vocabulary.
Common verbs (frequency of the verbs said, saw, went, was, were, got, get)	A high % indicates limited verb choice. A more adventurous choice of verbs helps to make writing more lively and interesting for the reader.
Use of adverbs to ascribe shades of meaning	A high % indicates a more sophisticated vocabulary.
Common words (verbs: said, saw, went, was, were, got, get + connectives and, so, but, then, because)	A high % indicates a limited vocabulary.
Commas	The use of commas to demarcate clauses is a sign of advanced writing skills. This is the ratio of average sentence length in words to the number of commas used (not including commas in lists). A high number of words to commas indicates a lack of commas. A low number of words to commas indicates mostly correct use of commas.

4.2.1 Ramayana Sub-Study

Control and experimental group first stories show four differences: the control group first stories score better for simple sentence starters, and the experimental group first stories score better for common verbs, common words and common connectives. Following the intervention, at the second-story stage all these differences disappear except common verb usage, which remains better in the experimental group. Additionally, the experimental group second stories use more words than the control group second stories, and both groups improved use of adverbs; the latter is consistent with the fact that their stories were comparable in this respect both before and after the intervention. These comparisons are summarised in the top half of Table 3; each group's average score is presented and the p-values for each t-test involved are given in parentheses. The significant second-story difference in the average total number of words used by the groups must be contrasted with the fact that no significant experimental group improvement effect is found, and this difference is therefore judged unlikely to be indicative of a strong effect. Finally, the bottom half of Table 3 show three weak degradations or improvements whose effect was not sufficient to result in any second-story difference, and two aspects in which no differences occurred.

To summarise, the strongest effects, that is, those for which an improvement or degradation is found which is consistent with the corresponding first-story and second-story differences, are:

- **A benefit to the experimental group** in use of simple sentence starters;
- **A penalty to the experimental group** in use of common words and common connectives;
- **A benefit to the control group and the experimental group** in use of adverbs.

4.2.2 Visitor Sub-Study

In this instance, both the first-story comparison and second-story comparison yield no statistically significant differences; in other words, the two groups' stories are comparable to each other, in the sense of every CHECK TEXT indicator, before and after the intervention. Longitudinally, the control group improved their scores for average sentence length (from 18.3 words to 10.5 words, $p=0.01$) and use of common words (from 14.1% to 10.7%, $p=0.01$), and both groups improved their scores for use of the word 'and' (control from 4.7% to 2.4%, $p=0.03$; experimental from 4.1% to 2.0%, $p=0.02$). Since these improvements are not reflected by any second-story differences, as with the Ramayana sub-study they are judged to be weaker effects. No degradations were found.

4.2.3 CHECK TEXT Analysis Summary

It is suggested that the main conclusion to be drawn from the above results is that no one intervention produced clearly different results from the others, in the sense of the CHECK TEXT indicators; it is particularly notable that the improvement effects in one experimental group do not show up in the other.

Further, the results of the Ramayana sub-study experimental group seem internally inconsistent. The improvement effect noted for use of adverbs and

Table 3

Ramayana Sub-Story: Cross-Sectional and Longitudinal Story Comparisons Using the CHECK TEXT Categories

	Simple sentence starters		Common verbs		Common words		Common connectives		Total number of words		Use of adverbs	
	Control	Experimental	Control	Experimental	Control	Experimental	Control	Experimental	Control	Experimental	Control	Experimental
First stories	20.7%	51.4%	6.4%	4.4%	12.2%	8.8%	1.7%	0.6%	116.8	146.9	0.7%	0.3%
	control better (p=0.00)		experimental better (p=0.03)		experimental better (p=0.00)		experimental better (p=0.04)		no difference (p=0.17)		no difference (p=0.43)	
Second stories	47.0%	29.7%	5.0%	3.9%	11.1%	11.3%	1.9%	2.0%	127.4	206.6	3.1%	3.6%
	no difference (p=0.11)		experimental better (p=0.05)		no difference (p=0.79)		no difference (p=0.78)		experimental better (p=0.05)		no difference (p=0.56)	
Improvement or degradation?	no change (p=0.07)	improvement (p=0.01)	improvement (p=0.03)	no change (p=0.49)	no change (p=0.17)	degradation (p=0.01)	no change (p=0.59)	degradation (p=0.00)	no change (p=0.62)	no change (p=0.07)	improvement (p=0.01)	improvement (p=0.00)
The connective 'and'												
	Control	Experimental	Control	Experimental	Control	Experimental	Control	Experimental	Control	Experimental	Control	Experimental
First stories	4.2%	3.8%	18.3%	19.3%	0.1	0.2	16.7	17.0	68.6%	65.6%		
	no difference (p=0.60)		no difference (p=0.57)		no difference (p=0.33)		no difference (p=0.90)		no difference (p=0.19)			
Second stories	4.2%	5.5%	14.1%	15.7%	0.1	0.3	19.2	17.9	62.3%	61.9%		
	no difference (p=0.91)		no difference (p=0.38)		no difference (p=0.07)		no difference (p=0.64)		no difference (p=0.91)			
Improvement or degradation?	no change (p=0.93)	degradation (p=0.02)	no change (p=0.07)	degradation (p=0.02)	no change (p=0.93)	improvement (p=0.03)	no change (p=0.38)	no change (p=0.56)	no change (p=0.08)	no change (p=0.37)		
Average sentence length												
	Control	Experimental	Control	Experimental	Control	Experimental	Control	Experimental	Control	Experimental	Control	Experimental
First stories	4.2%	3.8%	18.3%	19.3%	0.1	0.2	16.7	17.0	68.6%	65.6%		
	no difference (p=0.60)		no difference (p=0.57)		no difference (p=0.33)		no difference (p=0.90)		no difference (p=0.19)			
Second stories	4.2%	5.5%	14.1%	15.7%	0.1	0.3	19.2	17.9	62.3%	61.9%		
	no difference (p=0.91)		no difference (p=0.38)		no difference (p=0.07)		no difference (p=0.64)		no difference (p=0.91)			
Improvement or degradation?	no change (p=0.93)	degradation (p=0.02)	no change (p=0.07)	degradation (p=0.02)	no change (p=0.93)	improvement (p=0.03)	no change (p=0.38)	no change (p=0.56)	no change (p=0.08)	no change (p=0.37)		
Variety of words												
	Control	Experimental	Control	Experimental	Control	Experimental	Control	Experimental	Control	Experimental	Control	Experimental
First stories	4.2%	3.8%	18.3%	19.3%	0.1	0.2	16.7	17.0	68.6%	65.6%		
	no difference (p=0.60)		no difference (p=0.57)		no difference (p=0.33)		no difference (p=0.90)		no difference (p=0.19)			
Second stories	4.2%	5.5%	14.1%	15.7%	0.1	0.3	19.2	17.9	62.3%	61.9%		
	no difference (p=0.91)		no difference (p=0.38)		no difference (p=0.07)		no difference (p=0.64)		no difference (p=0.91)			
Improvement or degradation?	no change (p=0.93)	degradation (p=0.02)	no change (p=0.07)	degradation (p=0.02)	no change (p=0.93)	improvement (p=0.03)	no change (p=0.38)	no change (p=0.56)	no change (p=0.08)	no change (p=0.37)		

simple sentence starters might suggest that their second stories use a richer vocabulary, but on the other hand the degradation effect present in use of common words and words greater than five letters tells precisely the opposite story. An even more pronounced contradiction emerges between both sub-study experimental groups when one considers use of the word ‘and’; here they flatly contradict each other, one improving and the other degrading.

Thus it is suggested that the software intervention produced no consistent effect on the quality of the participants’ stories, as measured by CHECK TEXT, compared with the other interventions.

4.3 Teacher-Coded Analysis Findings

The only cross-sectional difference found to be significant at the 0.05 level arises from the Ramayana sub-study, in which the experimental group’s first stories and second stories score more highly for the number of strong descriptions of characters or of actions characters do (control first stories average: 2.1 descriptions, experimental first stories average: 4.6 descriptions, $p=0.03$; control second stories average: 0.7 descriptions, experimental second stories average: 3.4 descriptions, $p=0.01$). There is no significant corresponding improvement or degradation effect.

Longitudinally, two improvement effects emerge; the Ramayana sub-study experimental group are found to have improved their use of narrative point-of-view from the first to the second story (from 1.3 occurrences to 3.3 occurrences, $p=0.03$), although not sufficiently to distinguish this from the control group’s story in the second-story comparison. Both Visitor sub-study groups are found to have improved their character descriptions (control group from 0.9 descriptions to 4.3 descriptions, $p=0.02$, experimental group from 0.9 descriptions to 4.9 descriptions, $p=0.03$). As with the CHECK TEXT findings, it is concluded that there is no consistent effect across both sub-studies; again the improvements and differences found in one case are not repeated in the other.

5 Data Analysis 2: Cartoon Comparison

Attention now turns to an analysis of the Ramayana sub-study’s intervention. This is similar to the teacher-coded analysis in that it consists first of qualitative coding, followed by comparisons of difference using unpaired Student’s t-tests with equal or unequal variance as appropriate. Concerning population sizes, since this comparison relies only on the intervention-produced cartoons, it includes all the participants who produced a cartoon, irrespective of whether they also wrote a first or second story. This results in a control group population of seventeen, and an experimental group population of thirteen.

It is important to acknowledge that there is no baseline data set for this comparison. However, such a set is harder to define and acquire here than for storywriting: during the intervention the two groups by definition did different tasks and so can be argued to have been utilising different sets of skills; the experimental group’s task very likely helped ICT manipulation and adaptation skills come to the fore, whereas the control group’s task probably relied more on physical drafting and representation. Consequently, it is difficult to imagine what kind of baseline test would accurately capture the total set of skills

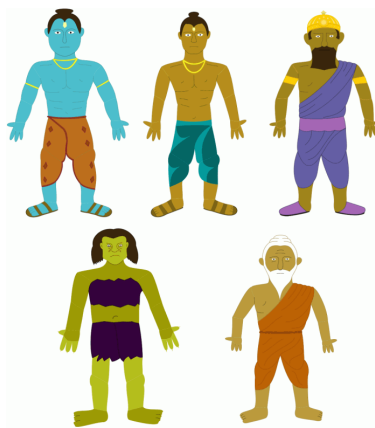


Figure 3: The characters provided to the experimental group for use within the software.

required over the two groups; a simple paper cartooning baseline would neglect prior ICT skills and could be argued to give one group extra practice, simply testing ICT skills is obviously insufficient and anything which tries to build a comprehensive picture risks becoming too complex and intrusive to normal class routine. Instead, it is intended that the class teacher's partitioning of the two groups, as described earlier, led to a reasonably equal distribution of skills between them. This certainly represents a comparatively quick and dirty approach, but in the context of the overall study it is judged acceptable.

Figures 3, 4 and 5 illustrate the complete set of resources available to the experimental group, consisting of the posable characters, some items of clipart and various backgrounds; some backgrounds were intended to be usable with no additional clipart, and others to be more customisable templates. From left to right and top to bottom of Figure 3, the characters are: Rama, his brother Lakshmana, their father, King Dasaratha, a female demon, and a holy man. The holy man asks King Dasaratha for his sons' help in vanquishing the demons who are bothering his group; the brothers travel with him to the forest where they fight one such demon. In addition to the front perspective shown, each character has a back view and can be made to open and close their mouth and eyes. The control group, in contrast, drew their characters on paper using colour as desired; they worked from the shared understanding of characters' appearances which the whole class had gained through earlier work on the Ramayana.

No attempt has been made to consider the role of the clipart provided within the software, and whether props and backgrounds are used or rendered differently on paper and in software, beyond the coding of how often characters are shown carrying objects. These questions are considered outside the scope of the current work; a proper study of them would likely require either a more comprehensive clipart bank or the facility for users to easily create their own artwork.

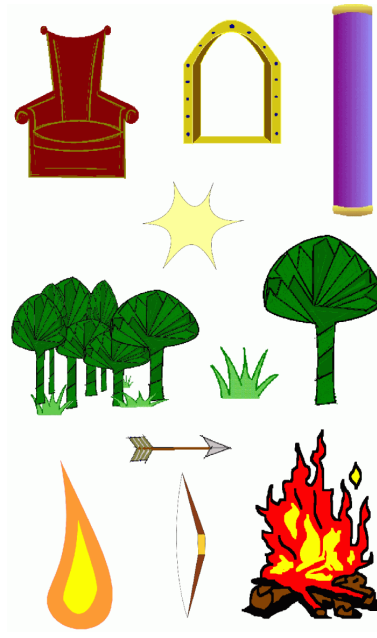


Figure 4: The clipart used by the experimental group.

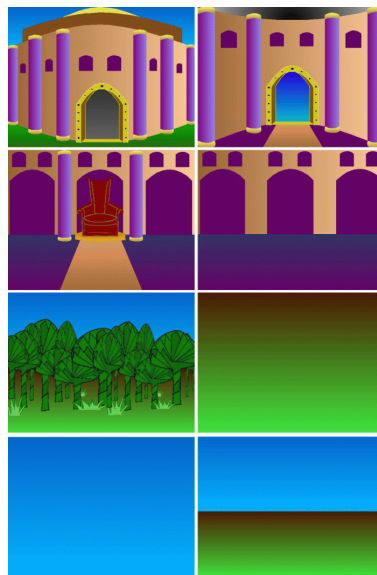


Figure 5: The backgrounds used by the experimental group.

5.1 Coding Scheme

The cartoons were analysed for three broad categories of feature, one aiming to detect any differences in use of camera language between the groups by examining the different shot distances used; and two focusing on characterisation, looking at facial expression and character pose. In each category, the intention is to get a measure of the degree of ‘expressiveness’ displayed by participants, loosely defined as involving any specific choice to depict something which deviates from a default option. This is easy to identify in the case of the software; characters’ appearances have a well-defined default state, and it turns out that such choices are usually obvious in the paper cartoons as well. Such an approach will camouflage instances in which a participant deliberately chooses default options for a picture, but this seems virtually unavoidable.

Expressiveness in shot distance, in characters’ facial expressions, and in characters’ poses will be considered separately, since these categories are fairly independent of each other. Comparisons are made between the control and experimental groups, and between the six male and seven female participants in the experimental group; for each category, two kinds of analysis will be used.

5.1.1 Fine-Grained: Differences in Specific Type of Expressiveness

This analysis examines the total number of times each participant makes a *specific* choice of signification from the options available in a given category, for example the number of times they specifically choose to deploy a close-up rather than another non-default shot distance. Each set of figures is averaged over the relevant control or experimental group to see if any significant differences arise in the groups’ specific choices of signification.

5.1.2 Coarse-Grained: ‘Expressiveness per Shot’

For a given category, this method measures the total number of times each participant makes *any* expressive choice and averages the total over the number of shots they make to give a rough sense of the participant’s ‘expressiveness per shot’. It is described as coarse-grained because it ignores the specific way in which the default has been modified. To see why this is necessary when averaging per shot, consider the alternative: that each participant’s use of every kind of expressive mode in a given category is averaged over their shots. Take as an example the depiction of emotion through facial expression; it is possible to imagine that each participant will then receive a figure for ‘anger per shot’, ‘happiness per shot’ and so on. The problem with this is that, in a visual story, there is not a clear relationship between the number of shots shown and the number of times a specific choice of depiction may be made; it is possible to choose to fill a sequence in many different ways. Some of these ways may include more depiction and emphasis of facial expression, others of another mode. Thus, for example, a longer sequence of shots themed on the same story topic as a shorter sequence *cannot* be expected to show simply a linearly increased level of some specific expressive choice, be it anger per shot or anything else; all that can be meaningfully asked is whether the total number of expressive choices in a given category has varied.

Table 4

Comparison of Shot Distances used in Cartoon Stories

	Control	Experimental	Girls	Boys
Long shots	0.6	1.2	1.0	1.2
	experimental greater ($p=0.04$)		no difference ($p=0.53$)	
Medium shots	0.4	0.7	0.9	0.5
	no difference ($p=0.50$)		no difference ($p=0.66$)	
Close-ups	0.1	0.3	0.1	0.3
	no difference ($p=0.36$)		no difference ($p=0.47$)	
Pictures made	5.5	9.5	10.6	8.3
	experimental greater ($p=0.00$)		no difference ($p=0.34$)	

5.2 Fine-Grained Analyses

5.2.1 Camera Language Findings

When coding for use of different shot distances, it is first necessary to clarify what the default option is and how to tell when it is used, since these concepts are not immediately obvious. The guide used is the finding of Burn and Parker (2003b) that children “showed a tendency to draw everything in long shot, as if needing to see whole figures against backgrounds all the time”; this sounds very much like a kind of default behaviour. Accordingly, a default shot distance will be chosen for each participant which is judged to most represent this description, and this shall be used for comparison with the rest of that participant’s pictures. Three non-default levels of shot distance are coded, corresponding to long shots, medium shots and close-ups; the non-default codings are only applied when a picture is judged to show a clear departure from the default mode identified above. The left-hand side of Table 4 summarises the results; it can be seen that experimental group participants produced on average about twice as many long shots as control group participants. No significant differences are found for medium shots or close-ups. Additionally, whether or not each participant produced a cartoon sequence which can be interpreted as finished, that is one which depicts events from the story’s beginning, middle stages and end, and the total number of cartoon frames s/he used to tell the story, is recorded. All but one of the thirteen experimental group participants finished their cartoons compared to fewer than half of the control group participants, and on average experimental group participants produced almost twice as many shots as control group participants over the two afternoons of work. As can be seen from the right-hand side of Table 4, no significant differences between boys and girls in the experimental group are found.

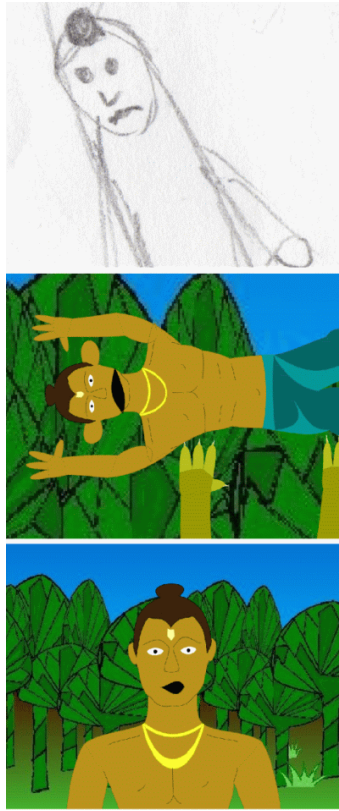


Figure 6: Example expressions included in the unhappiness/discomfort category.

5.2.2 Characterisation Findings: Facial Expression

Four different facial expressions are found in both groups' cartoons. These are happiness, unhappiness/discomfort, surprise and anger. The unhappiness/discomfort category is broader than the others, and in some cases it is arguable that a more specific description, such as fear, may be applied. However, such distinctions were not drawn so as to ensure a more robust classification. Figure 6 gives three examples of the variety of expressions included under unhappiness/discomfort; the intention is that this category does capture the broad intention behind each expression without trying to infer too much extra information. Examples of all four expressions may be found in Figure 7; the left-hand side shows control group images and right-hand side experimental group ones. Two numbers are given below each picture; the first shows each group's average number of depictions of the expression, and the second indicates how many participants in the relevant group depicted that expression.

Neutral expressions were also coded during analysis, though they are not included here because the concept of neutrality varies between the groups. Neutral expressions in the software cartoons are easy to define as the characters' default expressions, so that no attempt has been made to change any facial features. Although there might be instances where that default expression is considered

Table 5

Comparison of Emotions Depicted in Cartoon Stories

	Control	Experimental	Girls	Boys
Happiness	5.8	2.5	4.0	0.7
	control greater (p=0.03)		girls greater (p=0.05)	
Unhappiness/ discomfort	0.2	0.7	0.9	0.5
	no difference (p=0.23)		no difference (p=0.63)	
Surprise	0.4	0.2	0.3	0.2
	no difference (p=0.53)		no difference (p=0.74)	
Anger	1.2	0.8	0.3	1.3
	no difference (p=0.16)		boys greater (p=0.03)	

appropriate, it is judged that in the majority of software cases it has been used precisely as a default rather than as a deliberate choice. In the paper cartoons, some participants appear to have used a happy face as a default/neutral state (for example, when every character in almost every shot is smiling!), and others a more obviously neutral face. This variation between the groups seems to make the comparison of neutral faces largely pointless.

The total number of uses of a given expression was calculated for each participant, and these totals were averaged over the relevant group to yield the mean number of uses of that expression for that group.

Table 5 shows more than twice as many happy characters depicted, on average, using paper. No significant differences between the average number of depictions of unhappiness/discomfort, surprise or anger are found. The fact that happiness is the only area of difference in facial expression strengthens the above suggestion that many paper participants were tending to use happiness as a default expression.

The difference in happiness also carries over to gender, with the girls in the experimental group depicting almost six times as many happy faces as the boys. By contrast, boys depicted more than four times as many angry faces than girls on average. Both happiness and anger require the characters to be modified from their default expressions if the depiction is to count in the relevant total, so there may be a significant gender difference in character depiction here. No other gender differences are found.

5.2.3 Characterisation Findings: Characters' Poses

The range of character poses found is shown in Figure 8; most of the categories apply to single characters, but some encompass two, for example picking up another character, and some are combinable; for instance, sitting cross-legged whilst raising one's arms. The 'touching a character' category refers to nonvio-

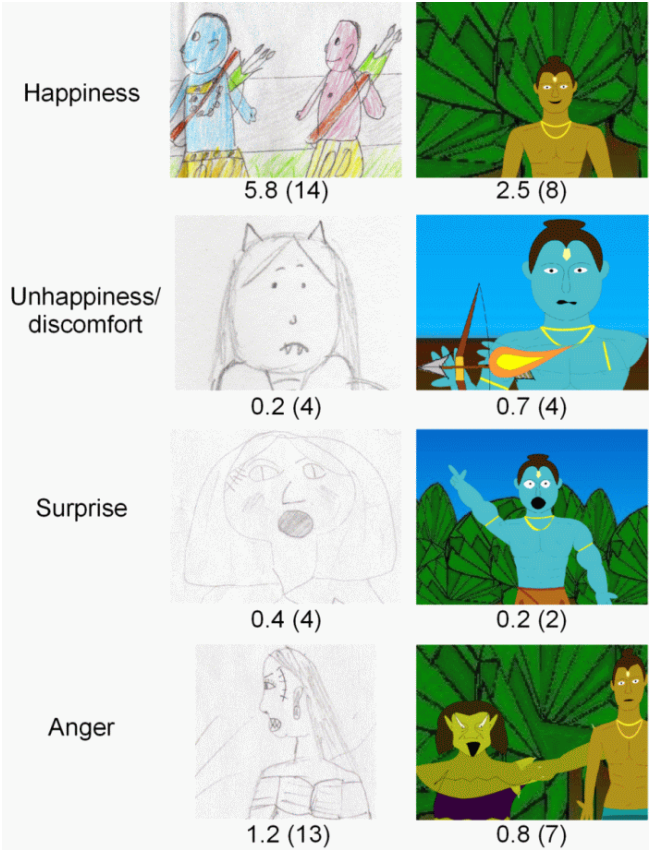


Figure 7: The four expressions depicted by both groups.

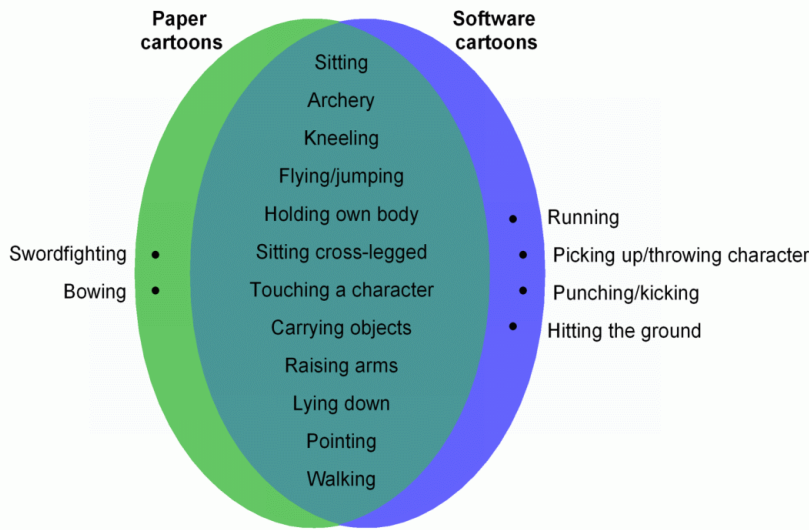


Figure 8: The sets of poses depicted in the paper and software cartoons, and their intersection.

lent touch; a separate count of ‘punching and kicking’ is made. An initial flavour of how poses look may be found in Figure 9, which shows a sample of the two groups’ renditions of a character holding their own body, and of a character sitting cross-legged. As with Figure 7, the control group and experimental group images are presented on the left- and right-hand-sides respectively, with the average number of depictions and number of participants making the depiction indicated below.

The counting and averaging proceeded in the same way as for facial expressions. Six statistically significant differences are found for the ‘universal’ poses which are common to both groups, in other words the poses in the intersection part of Figure 8; they are indicated graphically in Figure 10. It can be seen that the average control participant depicted characters kneeling more often (control average: 0.7 depictions, experimental average: 0.1 depictions, $p=0.00$), whereas the experimental group’s averages are greater for touching (control average: 0.1 depictions, experimental average: 1.3 depictions, $p=0.04$), flying/jumping (control average: 0.2 depictions, experimental average: 1.8 depictions, $p=0.01$), sitting (control average: 0.3 depictions, experimental average: 1.6 depictions, $p=0.00$), lying down (control average: 0.2 depictions, experimental average: 0.7 depictions, $p=0.02$), and raising arms (control average: 0.6 depictions, experimental average: 1.8 depictions, $p=0.02$). Figures 11 and 12 show what these poses actually looked like.

Figure 13 shows the average use of poses unique to one group or the other; here, the control group exhibits two unique kinds of pose and the software group shows four unique kinds of pose. Again, examples are provided in Figures 14 and 15.

Figures 10 and 13 suggest particularly that the software cartoons show much more physical interaction between characters; this is implied by the groups’

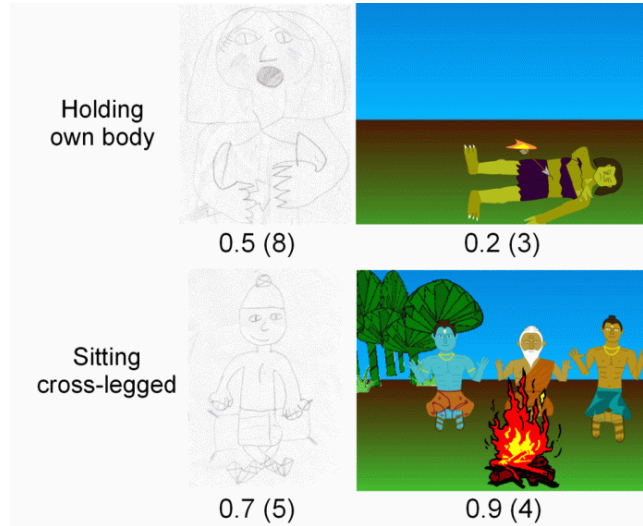


Figure 9: Examples of two poses depicted comparably often by both groups.

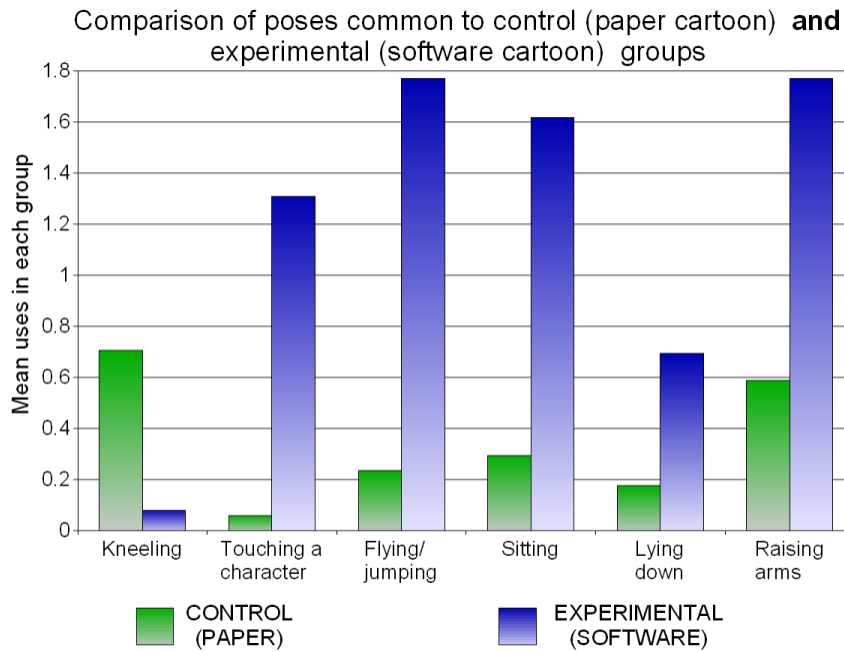


Figure 10: The statistically significant differences for the poses depicted by both groups.

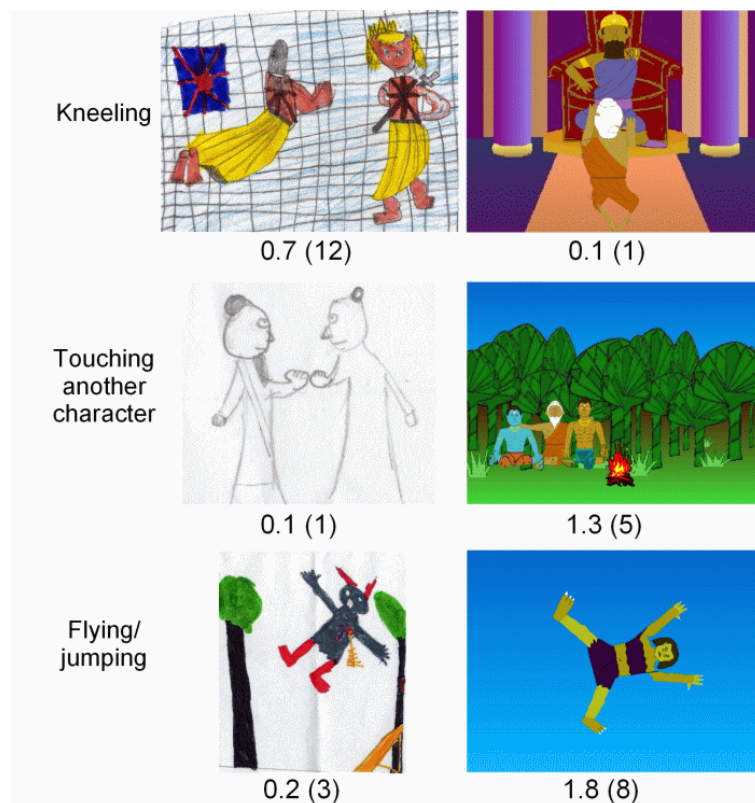


Figure 11: First set of example poses whose average use differs between control and experimental groups.

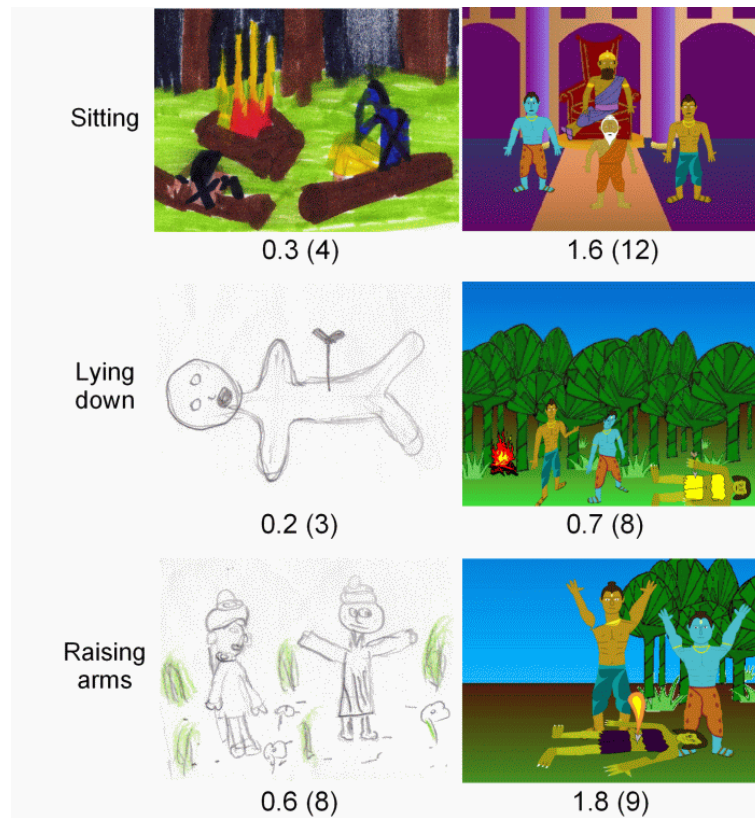


Figure 12: Second set of example poses whose average use differs between control and experimental groups.

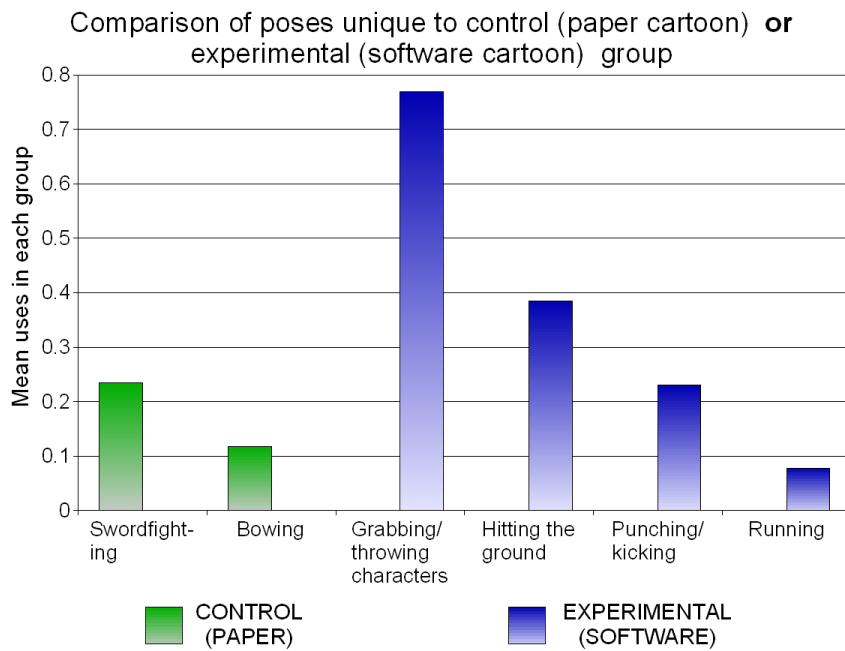


Figure 13: The statistically significant differences for the poses depicted only by one group or the other.

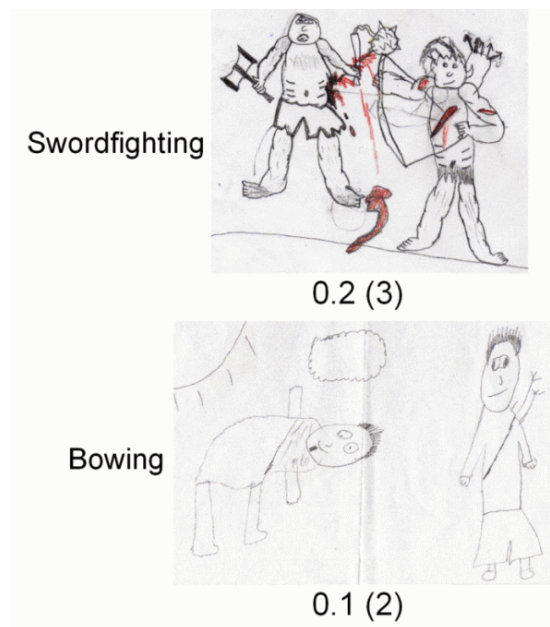


Figure 14: Examples of the two types of pose depicted only by the control group.

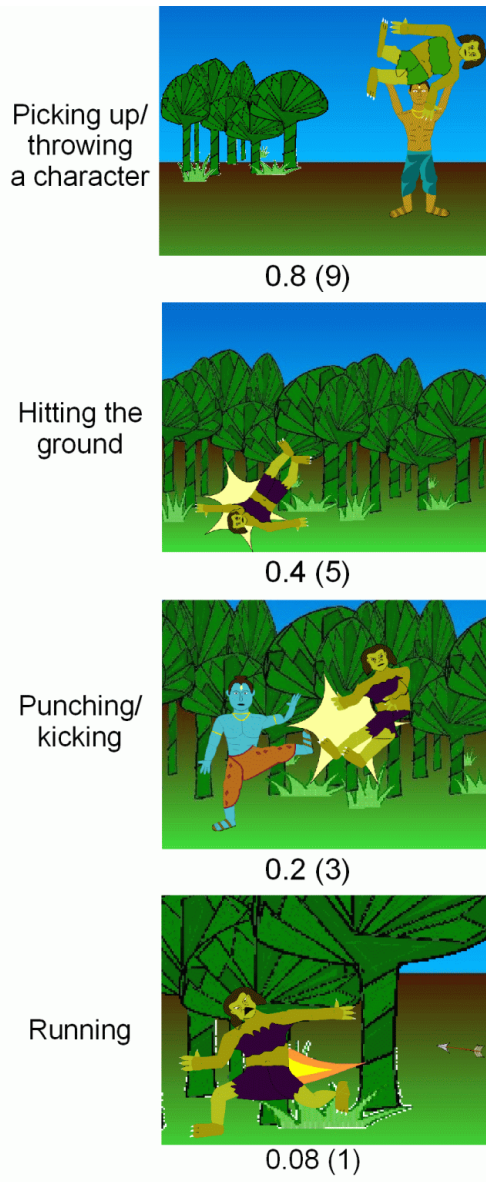


Figure 15: Examples of the four types of pose depicted only by the experimental group.



Figure 16: A pose which is arguably close to a bow.

differences concerning three types of physical interaction. First, the average number of instances of a character touching another is more than twenty times greater for the experimental group than for the control. Secondly, nine of the thirteen experimental participants showed a character picking up or throwing another character, and, thirdly, three showed a character punching or kicking another, whereas no control participant showed either of these actions. The experimental group's depictions of these three kinds of touch seem to be quite evenly split between the boys and the girls and no significant difference between their use of these pose types is found.

There is one kind of direct physical interaction between characters which is unique to the paper cartoons; three of the seventeen control participants depicted two characters swordfighting. The experimental group could not depict any swordplay, because their bank of clipart limited their choice of weapons to bows, arrows and fire. It is thus judged that the absence of swordfights does not weaken the trend of increased physical interaction in the software cartoons, but rather highlights an inflexibility of the software used in the study.

The other kind of pose unique to the control group's cartoons is that of a character bowing. It is possible that this pose, like that of kneeling, which was more common on paper, is more difficult to accomplish with the software, particularly since no side views of characters were provided; both of the bows were drawn from a straight side-on perspective. However, it is suggested that the pose may simply not have occurred to any experimental participant; one experimental group participant did create a pose which would be very like a bow if it were rotated, with the legs bent almost perpendicular to the body (see Figure 16). This, in conjunction with the fact that only two of the seventeen control participants thought of drawing the pose, is taken as supporting evidence that at least some experimental participants would have been able to achieve bows but that none thought of it.

As noted, the average control participant depicted kneeling characters more often, and this too may represent a type of pose which is currently difficult to render using the software. The single experimental participant who successfully depicted a kneeling character is also the only participant in either group to make a character run (see the bottom of Figure 15), so it seems likely that this participant was more adept than average at manipulating the characters and so was able to achieve comparatively difficult poses.

5.3 Coarse-Grained Analyses

Attention now turns briefly to the measures of expressiveness per shot for the three categories. Between the control and the experimental groups, the results show no significant difference in use of non-default shot distance per shot, that the control group averaged about three times as much facial expressiveness per shot as the experimental group (control average: 1.4, experimental average: 0.4, $p=0.00$), and that the experimental group may have averaged more pose expressiveness per shot (control average: 1.1, experimental average: 1.5, $p=0.55$). This might be interpreted in terms of what the software makes easier and what it inadvertently makes harder. It is suggested that, with the current version, facial expression is harder to manipulate than pose; it is very often necessary to zoom in so as to be able to drag a line, whereas a character's pose can be quite easily changed from a greater distance.

Within the experimental group, no gender differences in overall expressiveness per shot are found.

6 Qualitative Observations and Participant Opinions

This section will move from quantitative analysis to look at participants' opinions and experiences. Comments are drawn from the investigator's experiences during the intervention, an interview with the Visitor sub-study teacher, who had direct experience with using the software, and feedback forms filled in by the Visitor sub-study class after everyone had been given a chance to use the software.

6.1 Responses to Posable Characters

The ability to pose and warp characters was viewed both as one of the software's best features and one of its hardest; a few participants expressed both of these opinions simultaneously! Opinions as to best thing about the software include the following.

- “You can do whatever you want with the characters”;
- “Being able to move the skeleton and people”;
- “You can change the face's features”;
- “The best thing is that you can do a lot with the people”.

By contrast, responses as to what was most annoying or hard feature comments like

- “You couldn't do the right thing that you wanted to do”;
- “I could not get the skeleton to lie down”;
- “Moving the legs and arms”;
- “The little buttons to change the shape”.

This dual reaction suggests that the possible characters are seen as being valuable, challenging and problematic, and likely need user interface improvements.

There is a strong sense that users are keen to take more ownership of the artwork, and to adapt and individualise it; one feedback question asked whether users would like to be able to make their own characters, and many responded affirmatively to this, also asking for a greater variety of characters, and the ability to change their clothes. Further, during the intervention, a significant number of experimental participants were seen to make use of the ability to change the colours in the character drawings, lending support to the idea that allowing students to further customise artwork, or to create their own, would be a well received feature.

6.2 Use of Other Software Features

Participants did not easily make spontaneous use of close-ups, needing much encouragement to do so, as was also found by Burn and Parker (2003b). Participants also needed to be prompted to extend the initial short but complete sequences that most had completed by the end of their first session. These tended to consist of three or four sequenced pictures telling the bare rudiments of the story, after which a common reaction was to express the opinion that they had finished. The fact that virtually all went on to expand these starting sequences, after it had been made clear that this was the desire in the subsequent sessions, shows that a kind of editing behaviour was happening, although it was mostly limited to the act of ‘fleshing out’, inserting extra detail and extra shots, and did not extend to editing the sequence in which shots were presented. It is judged that the ability to display a slide show of one’s work in full-screen was valuable; participants made much use of this at the end of the intervention to show their work to peers and the teachers.

The visual storytelling prompts received very little use unless participants were ordered to use them; software logs show that the story summary window was occasionally requested, but the subsidiary ideas, character and camera helps were accessed very rarely indeed, and it is likely that the current iteration of the software makes this help too low-key. The motivation for providing such prompts is still valid, however; both teachers expressed the opinion that they were desirable and the lack of resequencing and use of close-ups described above shows that users do need explicit prompting if they are to deploy a fuller range of visual techniques.

Considering general ‘computer literacy’, although some parts of the user interface caused difficulties, overall the participants proved very able to learn the tool’s features and effect purposeful designs, and the speed with which the Visitor sub-study group did so was noted by their teacher. It seems likely that children are generally becoming ever more familiar with diverse software interfaces, tasks and capabilities.

6.3 Engagement and Motivation

Reactions observed during the interventions and user and teacher comments all suggest that participants found the software to be very engaging. The Visitor sub-study teacher observed that the software users had “put a lot more into planning [the story] than the children who used the paper and pencil... those

children had quite a few hours on it; they were completely absorbed in what they were doing”.

This teacher was additionally of the opinion that the software was useful because it provided an engaging additional way to mediate storytelling concepts, saying that

“the more ways that you can present something, the more chance you’ve got of hitting more children, of pressing their particular button, and it [the software] did it with that”.

This point is echoed in other comments made, in which it was said that the software provided another way for children to gather their thoughts and that it made them think through the stages of their story more carefully. This latter point perhaps ties in with the ‘fleshing out’ noted above; participants may have had to think more carefully about the stages when faced with the task of inserting more detail to their ‘first drafts’. This kind of behaviour is very likely easier in software, where the product is digital and can be edited and changed more easily than a paper plan or sequence can be.

Irrespective of how or why the software was beneficial to them, the participants themselves made many positive comments; as with the desire to be able to customise or create their own characters, there was a definite sense that extra features and improvements were wanted. A few of the comments made include “It was great it can be improved”, “It was brilliant but it can get better still”, “It’s one of the best things on the computer I have ever been on. It can still get better though”. The Visitor sub-study teacher also responded positively to the software, asking “when it’s done, can we have a copy?”

7 Summary and Conclusions

There may seem to be a conflict between the findings on storywriting and those on cartooning: the Visitor sub-study suggests that the software cartoon process had the same effect on writing as the conventional preparation process, and the Ramayana sub-study suggests that the software cartoon process also produced similar storywriting results to the paper cartoon process. However, the software cartoons were very different to the paper ones, showing more character action and interaction, and more shots.

It is concluded that these differences in cartoon depiction are real, but that they did not carry over to writing; given that arguably the greatest average difference between the software and paper cartoons, that is the number of shots produced, did not have a major effect, it may not be surprising that the other differences also failed to manifest themselves in print.

This cannot be attributed to participants being fazed by the notion of using pictures as a lead-in to writing, because both teachers confirmed that their classes were familiar with the idea, employing it in such activities as drawing cartoons of Greek fables and the Bible story of Exodus, or using images or sequences from live-action or animated versions of stories. The teachers also confirmed that the groups writing from software or paper cartoons did refer to their images as they wrote. In interpreting why the written stories failed to exhibit significant differences, then, attention must turn to the way in which the pictures were transcribed to text. The conclusion is drawn that participants did not possess picture-to-text transcription strategies capable of utilising the

greater degree of expression which the experimental group's cartoons demonstrated; although experimental group participants often made more expressive choices as to what to depict visually, it was not apparent to them how to translate these decisions into text. This perhaps demonstrates another area in which explicit process prompting may be valuable. Holdich and Chung (2003) demonstrate that software-mediated guidance in storywriting technique can lead to writing improvement, but the software used in this study did not provide explicit writing guidance, and this may be worth considering.

Despite this lack of impact on writing, it is concluded that the software led to a positive change in cartoon production and visualisation; software users produced more shots, were much more likely to finish their cartoon sequences, and showed the beginnings of greater expression through their increased use of long shots. In addition, it is concluded that the posable character facility helped software users express more types of pose, and to do so more frequently; the software cartoons showed more instances of character actions such as sitting, flying through the air, and lying down, and more character-character interaction, be it through depictions of touch, picking up, or fighting. Finally, the software participants were very engaged with their task, and all these effects together with the Visitor sub-study teacher's belief that the software users put more into their task than the others suggests that there is a definite productivity gain associated with using the software, at least as far producing visual stories is concerned; this may contribute guidance useful in the design of future tools in this area.

Additionally, taking this productivity gain with the lack of effect on writing suggests implications for other techniques, particularly those which seek to leverage students' facility or enthusiasm for certain artefacts and media. The role of computer games in education provides one example; even if a particular new process offers novel affordances for expressiveness and is received positively by students, care must be taken to see that these affordances can be transcribed to conventional products such as written work if they are to be seen to have strong educational value.

Some features of visual storytelling were either unaffected or made harder by the software. It seems that concepts of shot sequencing and juxtaposition were not employed any more fully than on paper, and they, like the transcription process itself, may need more explicit attention if they are to be of full value in literacy activities. Additionally, the current version of the software probably made the depiction of emotion harder than on paper, although it is only the average number of expressions per *picture* which is lower, not (happiness excepted) the average number of different kinds of expression per *participant*. It is therefore concluded that this difference indicates a weaker area of the software's interface.

Lastly, the lack of use of the visual storytelling prompts, particularly the subsidiary ideas, character and camera helps, is notable. Given the arguments presented above for having such prompts, and the proven success of the idea on which they are based (Holdich and Chung, 2003), it is desirable to determine whether such visual storytelling prompts can have a positive effect. In future studies, assuming that users can consistently be persuaded to pay attention to the prompts, it may be possible to determine their effect by using an approach analogous to the case studies reported by Holdich and Chung (2003); in these, two computer tools were used: the full system and the a cut-down version which

supplied a single guiding prompt and reply text box, but gave no other help. Therefore, the visual storytelling prompts might be tested by supplying two versions of the cartoon storytelling software; one would supply the full range of subsidiary helps covering ideas, characters and the camera, along with the main guiding prompts, the other might only supply a very general main guiding prompt. The resulting cartoon stories could then be analysed for differences. Additionally, the visual storytelling prompt concept has so far only been developed for a 'feedforward' stage; the actual HARRY system also provides a second set of feedback prompts, organised in the same sequence of story stages, aimed at encouraging revision and editing. Determining an analogous procedure for visual stories could be worthwhile research, insofar as it would bring the visual storytelling prompts closer to the original, proven, concept. This might be accomplished by having a checking system in place to detect if users have, for example, used a certain number of pictures in each story stage, have used particular types of camera shot such as close-ups of a particular character or object, and so on. It is not proposed to attempt the implementation of a general 'picture analyser', but rather to detect the use of specific characters, shot framings, and perhaps sequences and so on. For example, by tagging character and art files with useful semantic information, it might be possible to detect whether a user has shown a shot of a monster followed by a close-up of a character's face, as a simple action-reaction sequence.

Continuing to consider future work, two aspects of process design could benefit from further research. First, the task of designing a writing process which assists transcription may still be fruitfully pursued, so that users' pictures might elicit their writing more fully. One possible approach involves making captions of pictures as a first draft of writing, then having users rework their caption text when it has been removed from its pictures. Another idea involves using the pictures to stimulate various different categories of textual response, for example allowing students to brainstorm metaphors, similes, descriptions of action, and so on, before using these 'snippets' as starting points for text generation.

The process of visual story-making alone also merits closer attention. For example, it is desirable to find ways to better encourage students to deploy concepts of filming and editing, so that they make informed and deliberate choices as to shot framing, sequence and purpose; one possible way to do this might involve the visual storytelling prompts, if they could be suitably developed. Additionally, the concept of revision and editing of pictures and story sequences may prove fruitful, since users were not observed to revise or edit their existing pictures or sequences very much, preferring to make new pictures to add instead.

The motivational effects and positive responses noted earlier are acknowledged as likely, at least in part, to be due to a 'honeymoon' period which come from using any new and even vaguely interesting technique; thus it is important not to read too much meaning into them. Arguably though, some of the positive response stems from the opportunity the software offered participants to take ownership of the digital resources at hand; this was evidenced most obviously by the extent to which they made different expressive poses, but also by the way in which they were observed to create their own backgrounds by layering and scaling images, and to change the colour of characters' clothes and occasionally hair. Finally, therefore, it is greatly desired to increase the ownership which users may take of their visuals, particularly by enabling them to make their own posable characters.

References

- bfi Primary Education Working Group (2003), *Look Again! A Teaching Guide to Using Film and Television with Three- to Eleven-year-olds*. British Film Institute.
- Burn, A. and Leach, J. (2004), *ICT and Moving Image Literacy in English*. In Andrews, R. (Ed.) *The Impact of ICT on Literacy Education*, RoutledgeFalmer, London.
- Burn, A. and Parker, D. (2001), *Making Your Mark: Digital Inscription, Animation and a New Visual Semiotic*. *Education, Communication and Information (ECi)* 1 (2).
- Burn, A. and Parker, D. (2003a), *Analysing Media Texts*. Continuum, London and New York.
- Burn, A. and Parker, D. (2003b), *Tiger's Big Plan: Multimodality and the Moving Image*. In Jewitt, C. and Kress, G. (Eds.) *Multimodal Literacy*, Peter Lang, New York.
- Chun, D. M. and Plass, J. L. (1996), *Facilitating Reading Comprehension with Multimedia*. *System* 24 (4).
- Cunningham, K. (2000), *Integrating CALL into the Writing Curriculum*. *The Internet TESL Journal* 4 (5).
- Higgins, C. (2002), *Using Film Text to Support Reluctant Writers*. *English in Education* 36 (1).
- Holdich, C. E. and Chung, P. W. H. (2003), *A 'Computer Tutor' to Assist Children Develop Their Narrative Writing Skills: Conferencing with HARRY*. *International Journal of Human-Computer Studies* 59 (5).
- Holdich, C. E., Holdich, R. G. and Chung, P. W. H. (2002), *Assessing Aspects of Children's Written Grammar: Automating the Process*. *Computers and Education* 39 (1).
- Madden, M. (2007), *Cartoons Beyond Clipart: A Computer Tool for Storyboarding and Storywriting*. Unpublished PhD Thesis. Loughborough University.
- Marsh, J. and Thompson, P. (2001), *Parental Involvement in Literacy Development: Using Media Texts*. *Journal of Research in Reading* 24 (3).
- Mayo, E. (1960), *The Human Problems of an Industrial Civilization*. The Viking Press, New York.
- Mayo, E. (1975), *The Social Problems of an Industrial Civilization*. Routledge and Kegan Paul, London.
- McCloud, S. (1994), *Understanding Comics: The Invisible Art*. HarperPerennial, New York.

- Office for Standards in Education (Ofsted) (2005), Annual Report of Her Majesty's Chief Inspector of Schools 2004/05. The Stationery Office (TSO), London.
- Oldham, J. (1999), The Book of the Film: Enhancing Print Literacy at KS3. English in Education 33 (1).
- Parker, D. (1999), You've Read the Book, Now Make the Film: Moving Image Media, Print Literacy and Narrative. English in Education 33 (1).
- Parker, D. and Pearce, H. (2002), Story Shorts - Using Films to Teach Literacy. Department for Education and Skills (DfES), London.
- Robinson, M. (1997), Children Reading Print and Television. The Falmer Press, London and Washington DC.