Journal of Literacy and Technology
Volume 13, Number 3: December 2012
ISSN: 1535-0975

Evaluation of students' digital animated multimodal narratives and the identification of high-performing classrooms

.

Paul Chandler, Ph.D
Senior Lecturer in Education
Australian Catholic University
paul.chandler@acu.edu.au

Len Unsworth, Ph.D
Professor in Education
Griffith University
l.unsworth@griffith.edu.au

Annemaree O'Brien
Doctoral Candidate
University of New England (Australia)

Volume 13, Number 3: December 2012
ISSN: 1535-0975

Abstract

Contemporary approaches to literacy embrace digital and multimodal communication, and this is increasingly recognised in the syllabi prescribed by various education authorities across the world. Insufficient attention has been given to the evaluation of multimodal texts in ways which are semiotically grounded, accessible to the teacher and scalable to larger research studies. We present an evaluation instrument that addresses these requirements. The application of this instrument to 81 texts drawn from 17 classes has established the viability of the approach and allowed a subset of 'high achieving' classes to be identified. The derivation of the instrument is described in detail, the final form presented, evaluator guidelines elaborated, and the rating scales developed in full. Limitations are discussed along with recommendations for further work and development, but as an evaluation initiative the current work is presented as an important contribution to the continued development of multimodal pedagogy.

Volume 13, Number 3: December 2012
ISSN: 1535-0975

The reconceptualization of literacy in the context of our increasingly digital, multimodal information and communication world is now becoming more widely and prominently recognized in the curriculum requirements of government education authorities (England 1999; Australia 2009; Singapore 2010). With electronic, multimodal texts not only the dominant and preferred medium of today's digital generation, but also a required aspect of mandated curriculum documents, teachers in various ways have been weaving multimodal literacy into students' interpretive textual experience and to a somewhat lesser extent, into their text creation experience. However, while new digital multimodal literacies pedagogies are evolving (Anstey and Bull 2006; Unsworth 2008; Mills 2010; 2011), relatively little attention seems to be given to the development of an appropriate approach to the assessment of multimodal literacy development (Unsworth and Chan 2008; Unsworth and Chan 2009), especially in relation to students' creation of multimodal texts (Baxter and Simpson 2008; Kimber and Wyatt-Smith 2008; Macken-Horarik 2008). A very general conceptualization of an approach to evaluating dynamic digital and filmic texts produced by junior high school students in response to literature in the English classroom has been devised by Kimber and Wyatt-Smith (2008). They see the evaluation of such texts at the intersection of the textual evidence for students' *e-proficiency* (skill in the utilizing software affordances) and the quality of the textual *content*, *cohesion* and *design*. While this is a useful framework there is no specification of criterial textual features that would differentiate quality in respect of the four dimensions, and from the descriptive accounts of them it is difficult to clearly distinguish between *cohesion* and *design* features. On the other hand, working with primary school students' production of claymation, stop-motion movies, Mills (2011), following the work of Andrew Burn (Burn and Parker 2003; Burn and Leach 2004; Burn and Durran 2006), has related the meaning-making affordances of filmic text production to the "grammar of

Volume 13, Number 3: December 2012
ISSN: 1535-0975

visual design" extrapolated by Kress and van Leeuwen (2001/2006) from Hallidayan linguistics (Halliday 1978; Halliday and Matthiessen 2004). This enables a specific text-focussed, differentiated analysis of the students' movies providing the basis for feedback on learning and guidance to teaching. A somewhat similar approach was used by Thomas (2008) to discuss the quality of machinima produced by primary school students. However, this work has not extended to the formulation of a validated consistent procedure or specific instrument for systematically evaluating students' filmic texts.

This paper addresses the evaluation of digital animations in the context of our work facilitating middle years students' creation of 3D animated narratives (Chandler, O'Brien and Unsworth, 2010). Authoring narratives using 3D animation involves the students using computer software to create a movie in many respects similar to movies created using live-action work with a video camera. To write/create using 3D software involves harnessing systems of choices for making meaning. The complete repertoire of meaning-making resources available in 3D multimedia is quite simply vast, framed (for example) by Cope and Kalantzis (2009) in terms of five modes: linguistic, visual, spatial, gestural and audio. To support students' ongoing development of 3D animation authoring and to determine the most efficacious teaching practices, we need to derive a means of assessing the effectiveness of students' deployment of these multiple meaning-making resources. We have sought an instrument that teachers and researchers can apply systematically and relatively quickly in responding to students' work, providing informative feedback, and which could be scaled to provide systematic evaluation of several hundred texts. In this paper, we outline our approach which (a) attends carefully to intra-modal meaning (b) has the capacity to attend to inter-modal meaning (c) is suitable for the evaluation of relatively briefs texts developed by young, inexperienced authors and (d) can be readily applied to the bulk analysis of texts.

Volume 13, Number 3: December 2012
ISSN: 1535-0975

The centrepiece of this article is the development of the evaluation instrument itself, the use of which is demonstrated through the identification of 'high performing' classrooms. The identification of such classrooms is important so that subsequent work can draw on other observational and case study data in order to explicate features of teaching and learning which are important in the creation of high quality multimodal texts by school age students. The discussion proceeds as follows. Firstly, the framework for an evaluation instrument is described in principle, followed by a presentation of the particular evaluation instrument used in our investigation. The detail of how that instrument was used to broadly discriminate between the quality of work from 17 classes is then presented. The paper concludes with a discussion of the efficacy of the approach and implications for future development and application.

Evaluation by attending to semiotic systems

The starting point for the evaluative approach we are advancing is to consider the systems of choices that a creator of multimodal texts makes. A simple description for still images of the manner in which various design elements (or codes) and conventions together form a system from which combinatorial selections are made to convey meaning was provided by Anstey and Bull (2006, p. 108) and is shown in Table 1.

Table 1: Design elements and conventions in still images combine to make meaning (Anstey & Bull, 2006, p. 108)

The design elements of	*are combined through*	*the conventions of*	*to make meaning*
• Colour • Texture • Shape • Form		• Balance among design elements • Layout (how attention is attracted and focused) • Vectorality (how the eye is led through the image)	

We adapted this description to form a 'template statement' as the basis of our approach to

evaluation

> the text makes meaning by attending to *[category of meaning]* by strategic deployment of conventions associated with *[design element]*

For instance, "the text makes meaning by attending to *still images* by strategic deployment of conventions associated with *colour*", and separately, "the text makes meaning by attending to *still images* by strategic deployment of conventions associated with *texture*", and so on. Thus, an evaluation of a whole text can proceed by attending to relevant semiotic systems, and then to the codes within each system.

Using a range of such evaluative statements scopes the task of evaluation by identifying the semiotic systems that should be considered, and degree of delicacy to which they should be addressed. The form of the evaluation (e.g. written comment, yes/no, rating scale) is a separate consideration and our approach is addressed later in the section *Towards an Evaluation Instrument*. It should be observed that the template statement does not specifically include the conventions: it is assumed that the evaluator is sufficiently familiar with the relevant semiotics, although evaluator guidelines have been developed to facilitate confirmation of this for our purposes (see Appendix 1). An evaluator needs to work within the context of the social purpose of the text (in our case, a narrative piece, dealing with unusual or problematic events and their outcomes) and the socially constructed nature of conventions involved. For instance, black is the colour of death in some cultures, where as white carries that value in others; red conveys particular meanings in some cultures, but less so in others. Similarly, size, shape, proportions, clothing, hairstyle of characters will communicate important information. There are choices of colour, props, clothing and actual location which will situate the text in a particular time of day, season or era. Special effects such as fog can be used, and may variously signify a spooky environment or evening closing

Volume 13, Number 3: December 2012
ISSN: 1535-0975

in. In short, the approach we advance assumes that the evaluator is semiotically knowledgeable.

A list of evaluative statements is not intended to be either exclusive or definitive. The intention is to scope the evaluative effort – items 'to attend to' – not to provide a checklist that of items that must always be covered and never exceeded. For instance, by listing 'colour' as one of the evaluative statements does not imply that black-and-white line drawings would be automatically criticised – the evaluator can attend to the matter of colour and make a judgement that it is not relevant in this case. Rather, in thinking through how adequately the text addresses the meaning communicated through still images, the approach ensures that considerations of colour are not overlooked. Thus, the intention is that evaluative statements would help structure an overall evaluation of a text as an entire and coherent communicative enterprise. Later in the discussion, we will also note some examples of how authors/creators can attend to some of the design elements in particularly creative ways which belie treating the design elements or conventions as a mere checklist.

A further observation of this approach is that it tends to treat each semiotic system independently. If, for instance, a voice-over or background music were provided to accompany a still image then the design elements of each system would be treated separately, and there would be a risk of inter-semiotic meaning not being addressed. We need to identify this limitation from the outset and we have included a separate means for recognising those (see section on *Other Considerations*).

Having discussed, in general terms, an approach by which an attention to the multiple systems of meaning and the design elements thereof can be used to frame an evaluation, we now turn to the application of this approach to those systems of meaning available to the 3D multimodal author and readily used by young, inexperienced authors/creators.

Semiotic systems and codes identified from the software environment

Whilst this paper contributes to the general endeavour of improving the evaluation of multimodal digital texts, such texts can only be realized through the affordances of the particular (software) tools used to create them. The evaluation needs to take this into account, and is shaped, to some extent, by the functionality and capabilities of the software tool. Necessarily, we draw on an intimate knowledge of a particular item of software, *Kahootz* (Maggs, 2008), but in a way that would readily transfer to other products currently in the marketplace[2]. We also attend to the systems of choices that are relevant to the type of text regardless of the system used to produce it, which could also include techniques such as live action filming, stop-motion animation or claymation. Through close attention to the software functions available, two broad systems of choices can be identified which map onto two of the systems of meaning-making within Hallidayan linguistics (Halliday & Matthiessen, 2004): the ideational and interpersonal. The following presentation of the development of the template evaluation statement for the respective systems and codes has been informed by an insider's perspective of the software and references to key elaborations of the Hallidayan framework in relation to still image, moving image and sound (Kress & van Leeuwen, 2001/2006; van Leeuwen, 1996; van Leeuwen, 1999). The ensuing discussion concludes with a consideration of the third of the Hallidayan systems, textual meaning (or compositional meaning), and presentation of 'other considerations' which have been included in our instrument. That leads into a subsequent discussion of how these various components have been fashioned into a workable instrument by considering the issues of text-based evaluation, unit of analysis and a rating scale.

[2] Further examples of software of this type include *Muvizu* (http://www.muvizu.com/), *Kids Movie Creator* (http://www.kids3dmovie.com/en_01/Products.aspx), *Alice* and *Storytelling Alice* (http://www.alice.org), *Moviestorm* (http://www.moviestorm.co.uk/hub/australia), *Reallusion* (http://www.reallusion.com/) and *Anim8or* (http://www.anim8or.com).

Volume 13, Number 3: December 2012
ISSN: 1535-0975

'Creating a world': ideational and textual meanings

The representational/ideational system is concerned with communicating the nature of events, the objects and participants involved and the circumstances in which they occur (Unsworth, 2001, p. 18). As we proceed to elaborate, 'creating a world' principally involves these systems of meanings.

Using *Kahootz*, a 3D multimodal text is developed and presented as a series of scenes. The author selects one of many worlds (or sets) on which this scene is then further developed. It is not possible to import additional worlds, so one is constrained to work with a base palette from the library of worlds, but with the capability to re-colour or re-texture (i.e swatch) modify or appropriate them for a range of purposes. For instance, what initially appears as lush grasslands can be re-coloured to be a sparse desert. It is further possible to move through the world and thus choose a different location from the initially presented one.

Each scene can be populated with a range of objects, which can be selected from the extensive in-built library (as it the case with 'worlds', it is not possible to import characters into *Kahootz*). The object can be re-sized, have its proportions changed, and aspects of each object can be re-swatched. Thus, following the initial choices about setting and location, the author must choose how to populate the world, a task that embraces set dressing, props and characters. Included within this are decisions about how the objects are physically positioned, as it would be possible to have these (appear to) float in mid air or be (partially) buried in the ground. Furthermore, there are choices related to the arrangement of these objects – showing a group of characters who are looking at each other to represent a conversation, for instance. Objects can be animated – that is, caused to perform built-in actions or move from one location to another.

In addition to adding visual objects to a scene, audio can be added. *Kahootz*, for

instance, contained an extensive library of sound effects, along with the ability to record and import sound and together with the capacity for manipulation by specifying volume, pitch, echo, tremolo and duration. The audio mode, therefore, is a design element alongside choice of world, physical positioning, swatching, etc.

From the preceding discussion, it should be clear that there are three 'categories of meaning' which contribute to the ideational system:

- Setting and location (for instance, selection and swatching a set, identifying a location, adding various objects as set dressing, choosing lighting and special effects and including background sound effects or music)
- Participant selection and construction (for instance, selection, swatching and sizing participants and including dialogue)
- Arrangements and interaction of participants (for instance, the positioning of participants, the eyelines and gestures used)

These are presented in Table 2, along with design elements to which one might reasonably attend. Table 2 is intended to be read in the form of the evaluative template statement, for instance: "the text makes meaning by attending to *setting and location* by strategic deployment of conventions associated with *temporal location*".

The text makes meaning by attending to …	*by strategic deployment of conventions associated with …*
Setting and location	Choice of location
	Mood and atmosphere
	Temporal location (time of day, season, era, etc)
	Material location (environment, objects, inhabitants, etc)
	Material processes (animation of environment, objects, inhabitants)
	Narration &/or dialogue
	Sound effects

	Background music
Participant Selection/Construction	Material composition of a participant: Base choice of participant
	Material composition of a participant: Color of participant, clothes, etc
	Material composition of a participant: Relative size and shape
	Material composition of a participant: Animation
	Narration &/or dialogue
	Sound effects assigned to a character
	Background music assigned to a character
Arrangement and interaction of participants	Physical placement
	Physical arrangement
	Material processes of participants (interaction)
	Narration &/or dialogue
	Sound effects

Table 2: 'Categories of meaning' and 'design elements' for 'creating a world'

As foreshadowed above, Table 2 is not intended to be an exhaustive semiotic analysis, but in our experience describes the design elements associated with ideational meaning that may reasonably be deployed by younger multimodal authors using *Kahootz*. Table 2 does not prescribe that a multimodal text must have deployed particular coding systems or used particular conventions. Rather, it identifies, for instance, that meaning must be communicated about the 'setting and location' and in order to do so the author/creator should attend to a selection of the available coding systems which communicate meaning about 'choice of location', 'mood and atmosphere', 'temporal location' etc. Meaning-making in some of these categories is a necessary inclusion – for instance, a visual text must be located in some 'place' – whereas a clear indication of temporal location may not be important. Furthermore, some would be a necessary inclusion because a particular criterion was established for student work, such as a requirement that 'your text should be in the style of a murder-mystery set in the late 1800s'. In essence, the approach is designed to primarily reflect the Hallidayan systems of meaning in ways that guide an evaluator, but in a manner which is neither

Volume 13, Number 3: December 2012
ISSN: 1535-0975

constraining nor formulaic and is also flexible and extensible to accommodate variations in teaching and expectation.

Table 2 reflects certain limitations of both the software and the extent of educational experience of the students. For instance, *Kahootz* does not have capacity to lip-sync visuals with audio, and has almost no capacity for facial expression or for controlling the direction of gaze, while the capacity for gesture – particularly to convey emotion (e.g. anger, frustration, disinterest) – is extremely limited, and certain complex actions (such as the hand of a character to be realistically shown to clasp an object) are almost impossible to achieve. In the context in which we worked, the primary effort was effective visual communication, and the audio mode treated in much less depth. Therefore, the three systems of codes related to the audio mode (narration, sound effects, music) are described to a much lesser degree of delicacy than the visual codes, and an appraisal of inter-modal collaboration was not able to be dealt with in any significant fashion. Should a teaching sequence allow for a more detailed consideration of the auditory mode, then a greater degree of delicacy (e.g. pitch, pace, timbre) can be represented in an analogous manner to the visual codes. The inclusion of a separate category under which inter-modal collaboration is considered in the section on *Other Considerations*.

Whereas much of Table 2 is self-explanatory, some entries require explanation, such as how an author chooses, and works with, participants. In *Kahootz* (and the issues are much the same for similar items of software previously mentioned), material objects (people and things) are selected from a closed library where they are already coloured, clothed, have a defined shape and have a designated initial size. Within limits, the author/creator is able to develop variations of each of these aspects. So if a character or object doesn't quite 'work' within an animated text, it is possible that the issue is not that the author/creator has not given

Volume 13, Number 3: December 2012
ISSN: 1535-0975

sufficient attention to the design of the character, but the initial choice from the library (the 'base choice for the participant') was not ideal. For example, we have observed student work to retell the legend of Theseus and the Minotaur in which the student chose a bull as the base object for the minotaur. Despite attempted modifications, in the end, it still looked rather like mutant bovine. It is important, therefore, that the evaluation instrument take into account the 'basic selection' to the extent that it can be known. Similarly, there are all manner of problems which might beset authors/creators when actually placing the participant in the world, such as someone who is supposed to be standing on the ground appearing to be sunk into it (physical placement), people who are supposed to be standing close to one another actually appearing to overlap in space (physical arrangement) or someone who is supposed to walk towards another actually moves in an inappropriate direction (material process). These are all matters which one would presume the animator would have 'solved' in professionally-produced animations, but might still be features of texts created by young multimodal authors, and are therefore relevant considerations. We further observe that it is important to consider the participants (the living creatures who the narrative is 'about') to a greater degree of detail than either props (inanimate objects crucial for the story) or set dressing (objects chosen to elaborate the setting and location), and in this way the evaluation instrument implicitly elevates the importance of working with the participant in the overall communicative act.

Finally, it needs to be observed that 'creating the world' embraces textual meanings along with ideational meanings. Simplistically put, textual meaning embraces two concepts: firstly, the *modality*, or the extent to which the representation is naturalistic, realistic or minimally generic; and secondly, *composition*, dealing with how the various elements are integrated into a coherent whole. Kress and van Leeuwen (2001/2006) discuss in detail the

Volume 13, Number 3: December 2012
ISSN: 1535-0975

use of colour to portray modality and in particular colour attributes such as saturation, depth, illumination, brightness – attributes which are deliberate choices of the painter or photographer, but not necessarily available for the young multimodal author, given the state of development of the software. For instance, the deliberate choice of using a photo-realistic figure in a cartoony landscape 'says something' about the believability of either the set or participant. So, elements of modality are addressed in relation to both choices of colour and choices of participant; the examples of either the bull as the minotaur or the African animal in an Australian landscape can't be merely dismissed as inappropriate choices as they fundamentally contribute to the believability and modality of the text. This is reflected in the guidelines (Appendix 1), but also needs to embodied within an evaluator's appreciation of the design elements.

As a visual semiotic concept, composition is possibly easier to grasp. Firstly, it is suggested that there are socially-constructed conventions associated with how participants are placed with respect to one another, such as the placement of one participant to the left or right of another carries certain meaning, as is deliberately placing an object or participant in the centre or the periphery of the screen. There are also visual techniques which highlight certain participants rather than others (salience) and the use of devices to connect different elements together (such as dressing all members of a family in the same shirt), or showing a group of people as friends by locating them close to each other and facing each other. Therefore, issues of participant arrangement are not restricted to technical execution, and it emphasises that choices relating to colour and texture are multidimensional ones, as is reflected in the guidelines (Appendix 1).

Having developed an evaluation instrument which considers the communication of ideational and textual meanings, we now proceed to discuss the construction of inter-personal

meanings.

‘Showing a world’: interpersonal meanings

In addition to ‘creating the world’, the author is engaged with making choices about how that world is ‘shown to the viewer’. From the functional social scemiotic perspective (Halliday 1978; Kress and van Leeuwen 2001/2006; Halliday and Matthiessen 2004), it is through camera work and character positioning and movement that the nature of relationships between the participants in the text and the viewers of the text is established – the interactive/interpersonal systems of meaning (Unsworth 2001, p. 18). Through the camera, the multimodal author has ways of constructing social distance, social power and the extent to which the viewer is onlooker or participant, and through moving and relocating the camera, how these relationships change over the course of the text.

Five ‘categories of meaning’ can be identified through which the ‘viewing experience’ can be understood, arising from camera use, The categories of meaning relating to the viewing experience are:

- Sequencing of information (design elements which influence the order and the pace in which the information is presented)
- Viewer stance (that is, decisions relating to point of view)
- Camera distances (that is, to convey meaning related to social distance, and the use of the camera to hide or reveal information)
- The angles through which the information is seen (vertical camera angles conveying meaning related to social power, and horizontal camera angles conveying meaning related to involvement)
- The movement of the viewer with respect to that which is viewed

These are presented in Table 3, along with design elements to which one might reasonably attend.

The text makes meaning by attending to …	*by strategic deployment of conventions associated with …*
Sequencing of information	Sequencing of scenes
	Ordering of shots
	Pacing
	Narration &/or dialogue
	Sound effects
Viewer stance	Point of view
The selection of visual information (framing)	Camera distance
	Hide/Reveal
The angles through which visual information is seen	Vertical camera angles
	Horizontal camera angles
The movement of the viewer with respect to that which is viewed	Camera movement (e.g. zoom, tilt, pan, track and fly)

Table 3: 'Categories of meaning' and 'design elements' for 'showing a world'

As for Table 2, the systems and codes described in Table 3 are not semiotically exhaustive, but many of them are essential considerations. For instance, the camera is necessarily always positioned at a particular distance and angle. Though slightly awkward, terminology such as "the angles through which visual information is seen" is deliberately used to emphasize that it is how the viewer perceives the visuals that is ultimately significant – it does not matter whether the creator/author has achieved this effect by moving the camera or moving the participants, or produces some visual illusion which has the same effect. Whilst there may be a most obvious construction technique, in nearly every case, there are multiple ways of attending to each design element. These systems of meaning are potentially even more inter-related and subtle than 'creating the world'. For instance, one might elevate

Volume 13, Number 3: December 2012
ISSN: 1535-0975

the camera to glimpse a train rushing towards the participants from a distant location, but to do so could simultaneously imply a change of power relationship with a participant on whom the camera was previously directed.

To summarise the discussion so far: a template statement which can be used to focus an evaluator's work on a particular systems of meaning has been presented, and eight categories of meaning which are readily realised in animation software such as *Kahootz* have been presented. Along with these categories of meaning, some 31 design elements have been presented, each of which contributes to the meaning making in a category. Whilst attending to all the design elements listed is not essential, each of these eight categories of meaning must be considered in an evaluation process. For instance, it is not possible to construct a text which is not located somewhere (though it may be in a somewhat nondescript locality, era or time of day) and the camera must necessarily be operating at a particular distance and angle. No active decision in relation to 'viewer stance' may well result in a text which is seen through the eyes of a distanced, dispassionate observer, but to the viewer this is nevertheless important. Likewise, a text viewed via a stationary camera may be interpreted by a viewer as more like a news report than an action/drama, and so making no active decision has important consequences for how the text will be viewed. Similarly, arguments for the essential nature of the remaining categories of meaning can be advanced.

Other considerations

The discussion above, particularly as summarised in Tables 2 and 3, fully describes the categories, semiotic systems and codes that form the basis of our evaluation of student-created 3D multimodal texts. There are four additional categories, which we have identified as being important to evaluating such student work:

Volume 13, Number 3: December 2012
ISSN: 1535-0975

- Our particular interest is with multimodal narratives. Therefore, we include the need for an evaluation of the quality of the orientation, complication, evaluation and resolution of the text, as well as an item simply asking ‘is sufficient information communicated to tell a story’. That is, one essential category of meaning and five essential design elements.
- On-screen text has not been considered in the preceding discussion, although it could have been included in ‘creating a world’. The use of such techniques for titles and credits sets it apart from other elements, and if used for any other purpose (e.g. the words ‘one hour later’ appearing to indicate the passage of time) the conventions are somewhat unclear, and we therefore make a judgement about these separately. That is, one optional category of meaning and two optional design elements (‘titles and credits’ and ‘on-screen text’).
- It is important to recognise that sometimes significant technical innovation is employed. This may only be noticed by a very skilled observer who is highly familiar with the software platform. We believe that it is important to have a category to recognise significant student effort where it contributes to the overall communicative enterprise. That is, one optional category.
- The approach taken tends to regard the codes as operating somewhat independently rather than interactively within each semiotic system. This is compounded because the degree of delicacy is quite different for the auditory and visual modes. We have sought to ameliorate some of these difficulties by including a category on ‘multimodality’ which permits a global judgement to be made about the collaboration between the aural, verbal and visual modes. That is, one optional category.

In summary, we have identified eight categories of meaning and four other categories which

are the framework for an evaluation of multimodal texts, which together comprise a total of 40 possible design elements. We now proceed to describing how this has been further fashioned into a workable instrument by considering text-based evaluation, unit of analysis and a rating scale.

Towards an evaluation instrument

Text-based evaluation

The approach we advance is to identify the 'literal' or 'concrete observable' elements present in the text. Rather than trying to infer the author's intention our approach is to direct an evaluator's thinking to the appropriateness of "what is actually there". For example, the text may show a person: What colour is that person's clothes or hair? Is that person animated? Does s/he speak? Are we seeing that person from a high angle or a low angle? Are there other people whom we do not see initially, but are later revealed to us? And, most importantly, what meaning is created for us through this assemblage of attributes in their combination?

In addition to a 'literal' approach, our approach is that of an etic, or outsider's, perspective. As Harris (1979, p. 32) notes, an emic (or insider's) perspective would suggest that the insider (in our case, the student who created the text) would be the ultimate judge of the observer's descriptions or analyses, whereas

> etic operations have as their hallmark the elevation of observers to the status of judges of the categories and concepts used in descriptions and analyses … Rather than employ concepts that are necessarily real, meaningful and appropriate from the native point of view, the observer is free to use alien categories and rules … Frequently, etic operations

> involve the measurement and juxtaposition of activities and events that native informants may find inappropriate or meaningless. (p. 32)

Therefore, establishing a theoretical framework at all, necessarily positions our approach as etic. Moreover, whilst we support the idea of such a framework to be used in self-assessment or peer-assessment approaches, our evaluation efforts (as described later) rely on the ability of a dispassionate 'outsider' to provide an assessment of the work.

Unit of analysis

The work produced by students in this study is of quite brief duration (frequently less than 30 seconds), with few, if any, changes of scene. For our purpose, the unit of analysis is the entire text. Therefore, there is a single global judgement made in relation to each evaluative statement. Since the quality of the work may vary across the text, the rating scale takes into account the possibility of variability, as is now discussed.

Rating scale for design elements

The previous discussion has identified that there are 40 design elements which may be considered. Essentially, we seek a global indication of how frequently in the work has appropriate use of each design element been observed. The starting point is a consideration of how frequently the design element has been deployed in the text, and then to consider whether that use is generally appropriate or is given an 'appropriateness' rating. For reasons of efficiency, a simple scale of 'high', 'medium' and 'low' was chosen as a basis for the rating (elaborated below). When we take into account the realities of work by young, inexperienced authors, our experience has suggested that two additional degrees of delicacy be added, concerning 'incompleteness' and 'distraction' which are described in due course.

This leaves open the question of how to make a judgement about 'quality work', or the

basis by which an evaluator may distinguish between ‘high’, ‘medium’ or ‘low’ for any design element. This is a vital consideration. In principle, it would be possible to establish a rubric by which an evaluator might work. With each design element, and the conventions applicable, being logically discrete this implies not one but up to 40 rubrics, the presentation (let alone development) of such a tool would mitigate against the practical usability and relative expedience of the evaluation. This would be made even more complex when one considers that there may reasonably be different expectations for shorter texts compared with longer ones and the fact of design elements combining to create meaning overall. The functional social semiotic perspective derived from Hallidayan linguistics provides for a rich rather than a reductionist understanding of text. We thus made the decision not to scaffold the evaluation in extreme detail. Rather, we make the assumption that the evaluator is familiar with the relevant semiotics and works under the general brief to look for meaning being made in each of the eight categories of meaning (Tables 2 and 3). The desirable 'frequency' with which a design element might be presented in a text will vary considerably from one design element to another. For instance, it may be quite effective for all camera distances to be close-ups. It is less likely that if all of the camera work were mid-shots that this would be effective, but it might take only one sensitively constructed close-up amidst an extended range of mid-shots for the camera distances to be rated as ‘high’. Thus we leave the assignment of ‘high’ ‘medium’ or ‘low’ to the professional judgement of the evaluator supported by the assessor guidelines (appendix 2).

Refining the principal consideration of ‘appropriate use’, we found it necessary to introduce a degree of delicacy to consider “To what extent should more have been done?” - the ‘incompleteness’ rating. The rationale for this is that student work may actually be incompletely thought through or implemented. For instance, a ‘foggy night’ might be

mentioned in narrative or voice-over, but not attempted visually; or there might be careful attention to detail of light/sun colour and direction in one time and no apparent concern for this later. Of course, both of these could be a deliberate choice of how to deploy the semiotic systems available, but it is more likely to be indicative of incomplete work arising a lack of time, attention or review. It is also a recognition that the unit of analysis being the whole text is very broad and that a design element might be quite well executed overall, but elements of incompleteness or inconsistency would still be present. Therefore, an incompleteness rating of 'high', 'medium' or 'low' or 'not applicable' is introduced. As is shown in table 4, the four-point 'incompleteness' rating subdivides the three-point 'appropriateness' scale to create a 12-point scale.

Appropriateness	L				M				H			
Incompleteness	H	M	L	NA	H	M	L	NA	H	M	L	NA
Overall score	1	2	3	4	5	6	7	8	9	10	11	12

Table 4: A basic rating scale for each design element in use

Further, we found it necessary to introduce a second level of delicacy to consider "To what extent has distraction been observed?" - the 'distraction' rating. This was included to take into account what we had already observed, informally, in the texts produced by young, inexperienced authors. Consider, for instance, a clown appearing unbidden into a desert scene, most likely because the student was experimenting with the clown and forgot to delete him/her, or because they thought it was funny or interesting (in the way that only an 11-year-old can!). Alternatively, a hippopotamus seen amongst the Australian animals in an Australian outback scene in the retelling of an indigenous story would be considered a distraction unless the hippopotamus was part of the story. It seemed to us that in order to take

into account the realities of student-developed texts, ‘incompleteness’ would sometimes be an insufficiently strong criticism, and that even with ‘incompleteness’ identified the text would (by definition) still be rated in the same main band. With these considerations in mind, we introduced a four-point judgement (‘high’, ‘medium’ or ‘low’ or ‘not applicable’) relating to ‘distraction’, which we see as a way of strengthening the ‘inappropriateness’ rating when necessary, and indeed the two are very closely related. We structured the rating scheme so that the ‘distraction’ rating could not be higher than the ‘incompleteness’ rating. In other words, a text could have low levels of distraction as a component of a modest level of completeness, but high levels of distraction can not be logically associated with low levels of completeness. The full rating scale is presented in Table 5. It both modifies the basic (appropriateness) rating and effectively adds a 'very low' band (score of zero) to the initial three-band scale. To clarify the earlier example: where only a small number of sound effects are used, but each of them is a poor choice, the basic rating must be a 'low', but the application of the distraction rating places them in the 'very low' band.

Appropriateness rating		L				M				H			
Incompleteness rating		H	M	L	NA	H	M	L	NA	H	M	L	NA
Distraction Rating	NA	1	2	3	4	5	6	7	8	9	10	11	12
	L	0	1	2		4	5	6		8	9	10	
	M	0	0			3	4			7	8		
	H	0				2				6			

Table 5: The complete rating matrix
illustrating how ‘appropriateness’, ‘incompleteness’ and ‘distraction’ combine to generate an overall rating; a blank is an invalid combination

In summary, the broad rating scale is a simple ‘high’, ‘medium’ and ‘low’ one, which has been elaborated to take into account the practical realities of working with young,

Volume 13, Number 3: December 2012
ISSN: 1535-0975

inexperienced authors/creators along with some inherent difficulties with treating the whole text (albeit a short one) as the unit of analysis. There is a strong hierarchy in applying the rating scales, which is reflected in the way the numerical ranking has been assigned (see tables 4 and 5). In particular, an evaluator will:

1. First consider "is consideration of this *[the design element]* a necessary inclusion?" If it is not, and there is little evidence of it, decide not to include it in an overall judgement of the design element.
2. Where it is a necessary inclusion, consider "how appropriate was the deployment of *[the design element]* within the category being considered?" This is the foremost consideration, and provides the basis for an 'appropriateness' rating. It is an 'impressionistic weighted average' of the semiotically appropriate use of the design element.
3. Then ask her/himself "is there much scope or necessity for doing more with *[the design element]*?". This is the basis for an 'incompleteness' rating.
4. And finally as him/herself "was there anything that distracted me in how *[the design element]* was exemplified?" This is the basis for a 'distraction' rating.

The highest possible rating on a design element is for one that has been used extensively, is used for good semiotic effect and there are no obvious instances of needing to 'do more' or of 'distraction'. Colour and lighting is an inevitable inclusion, and would probably be rated at the highest level as long as there was some variation to capture changing mood or the passage of time. The lowest rating is for a design element which has been used very little, but is highly distracting, for which an obvious example would be where sound effects are rarely used but are always inappropriate sound effects. Items which might fall mid-range on the scale would be a playground scene which is sparsely populated by people, playground

equipment, flora or fauna, but those few which have been included are quite suitable. Thus we have a means of providing an impressionistic rating for all design elements which is useful despite the different ways in which the design elements interact with construction of the text as a whole. The method assumes that the evaluator is familiar with the design elements and relevant conventions, but guidelines are also provided (Appendix 1).

Combining individual ratings into a rating for each category of meaning

The objective is to arrive at a global judgement of the effectiveness of each category of meaning. The above rating scale allows for a rating of each of the design elements in use, and these need to be combined into an overall rating for the respective semiotic system. This is done on the basis of an arithmetic mean of the rating derived from table 5. This is illustrated in tables 6 and 7, which consider the category of 'setting and location'. Table 6 illustrates how this is achieved when the evaluator has determined that all design elements should be considered, and Table 7 illustrates how this is achieved when the evaluator has determined that only certain design elements are relevant.

Design element	Appropriateness	Incompleteness	Distraction	Score
Choice of location	H	N	N	12
Mood and atmosphere	H	N	N	12
Temporal location: time of day, season, era, etc	M	M	N	6
Material location: environment, objects, inhabitants	H	L	N	11
Material processes: animation of elements	M	M	L	5
Narration &/or dialogue	L	N	N	4
Sound effects	M	H	L	4
Background music	L	L	L	2
Overall rating = Average of codes present = (12+12+6+11+5+4+4+2)/8 = 7				

Table 6: Aggregating scores – all design elements relevant

Design element	Appropriateness	Incompleteness	Distraction	Score

Choice of location	H	N	N	12
Mood and atmosphere	H	N	N	12
Temporal location: time of day, season, era				
Material location: environment, objects, inhabitants	H	L	N	11
Material processes: animation of elements	M	M	L	5
Narration &/or dialogue				
Sound effects				
Background music				
Overall rating = Average of codes present = (12+12+11+5)/4 = 10				

Table 7: Aggregating scores – sub-set of design elements relevant

The approach of using a non-weighted average of each of the design elements in use is consistent with the 'etic' and 'broad-brush' approach identified above. It is assumed that each design element in use is of equal importance to the meaning being conveyed through the particular semiotic system. Any discussion of relative importance would need to be developed separately as a close-up qualitative discussion of the text. Furthermore, it does not permit any consideration of what an evaluator might think that could have been used – the text stands on face value. This is one of the reasons for including the category of 'is sufficient information communicated to tell a story'. Consider the case of portraying a desolate outback scene: maybe the author/creator meant to show tumble-weeds rolling across the scene, but found it too difficult or time-consuming to do, meant to have a sound effect but encountered technical difficulties, or maybe never actually through at that level of detail. An evaluator taking an etic stance would neither have access to this information nor pay any heed to it – s/he would only be able to make the judgement that there were no sound effects (therefore rated as 'not applicable') and 'more could have been done' with the material location (therefore probably rated as 'high').

Volume 13, Number 3: December 2012
ISSN: 1535-0975

Interpreting numerical results

As it should be clear, an instrument developed on the above rating scales is capable of generating a vast range of numbers. Whilst some categories (e.g. 'point of view') consist of only one item and thus a score which is a whole number between 0 and 12 will be generated, other categories (e.g. 'location and setting') consist of several items, and thus generating a score which may subdivide the 12-point scale into 60 or more subdivisions. Such finesse is inconsistent to the general broad-brush approach, but equally rounding to the nearest whole number is an unnecessarily crude approximation and would mitigate against any assessment needing to pay attention to any more than one or two design elements in each category. In light of these considerations, we have made an arbitrary determination that the values for the rating of each category of meaning will be rounded to the nearest 0.25.

Implementation

For convenience, the evaluation instrument was set up using Microsoft Excel (Figure 1). In this way, an entry was only required for each of the columns ('should it be considered', 'appropriateness', 'incompleteness' and distraction) for each design element, and the correct value (Table 5) calculated automatically. The validation rule that the 'distraction' rating could not be higher than the 'incompleteness' rating was also automatically implemented, along with the calculations ensuring that only the design elements in use were included in the computation of a total score (c.f. Tables 6 and 7). Some time-saving strategies were also included, such as if there was no entry for 'incompleteness' or 'distraction' it was assumed to be 'not applicable'.

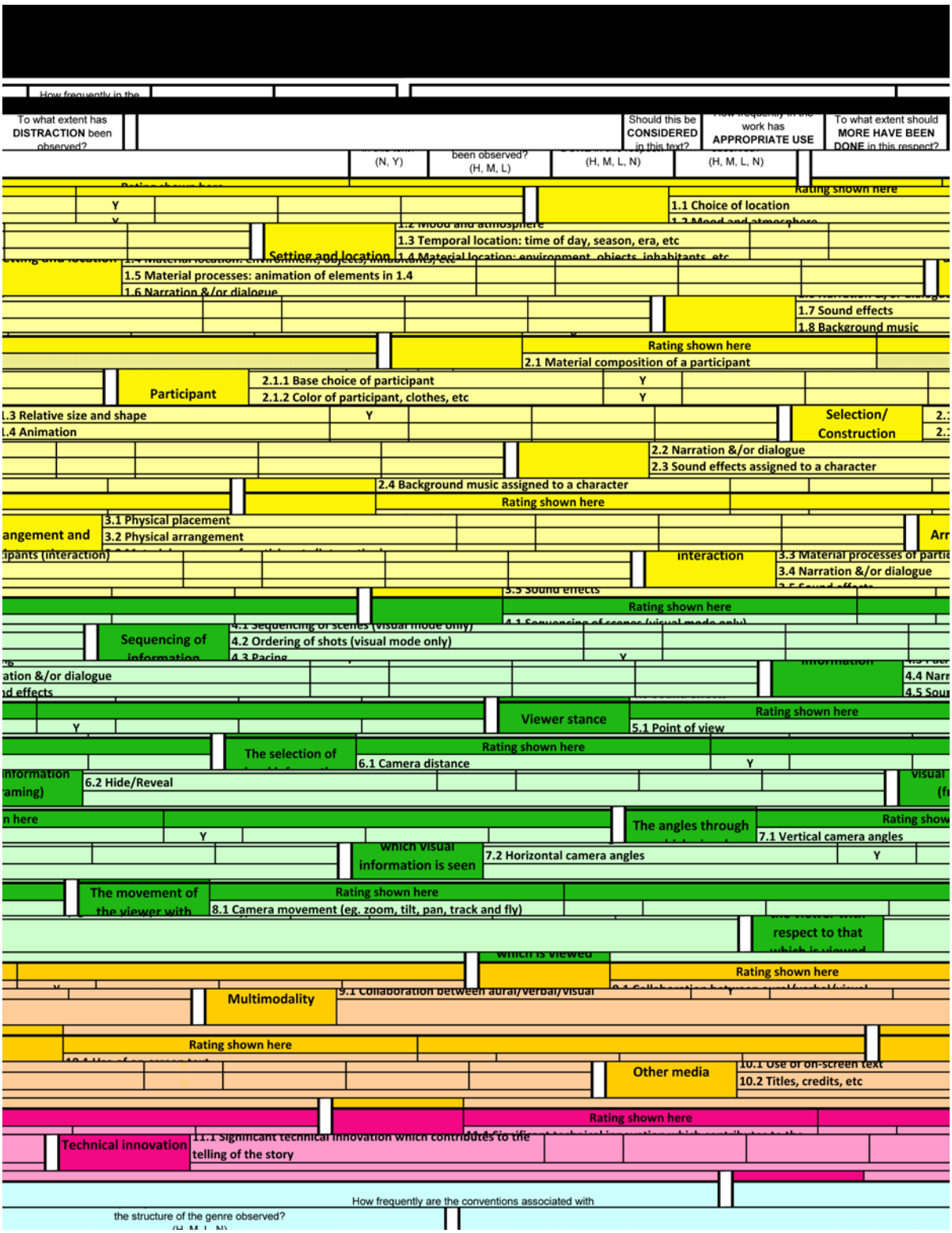


Figure 1: The Evaluation Instrument

Volume 13, Number 3: December 2012
ISSN: 1535-0975

The study

Participants

Over a period of a school year, we have worked with 17 upper primary school classes (i.e. children mainly between the ages of 10 and 12), their teachers and the students. School A (4 composite year 5/6 classes) was a government school in rural Victoria; School B (3 year 5 and 3 year 6 classes) was a government school in metropolitan Melbourne, Victoria; School C (one year 6 class) was a government school from semi-rural Tasmania; School D (1 year 5 class and a composite year 5/6 class[3]) was a government school in metropolitan Melbourne, Victoria; and School E (two year 5 and two year 6 classes) was a Catholic school in metropolitan Melbourne, Victoria. These teachers had volunteered to participate in a year-long program of introducing 3D multimodal authoring to their classes. The maximum number of multimodal texts which could be collected was 350.

Selection of texts

Teachers of each of the 16 classes were asked to identify the five best texts in their class, using professional judgement and broadly based on the student's attention to multimodal design. They were also asked not to be too pernickety about the selection. For instance, if they could not distinguish between the top seven (and from then there seemed to be a decline in quality) to select all 7. In total, 81 texts were identified by their teachers as the texts of highest quality, as described in Figure 2.

Evaluation and moderation processes

Two experienced teachers were employed to evaluate this corpus, and were provided with three hours of initial training and continual access to the researchers for advice. Texts were

[3] A team-teaching arrangement meant that both classes were, in essence, taught by the same teacher, so we refer to '16 teachers' throughout.

allocated to each review in a stratified random fashion, such that both evaluators saw examples from each class. A strict regime was established to ensure comparability in the work of both evaluators:

- work through the body of work 10 texts at a time
- each evaluator chooses a group of texts 10 texts from those they have been allocated
- each evaluator independently assesses the texts in the batch of 10 as per the evaluation instrument
- each evaluator chooses a "high", "medium" and "low" text for the two major evaluation categories 'elements present and observable' and 'elements as shown to the viewer'[4]
- this moderation sample is provided to the other evaluator for blind assessment
- the results from the two evaluators are compared for each text; to be 'sufficiently similar' the evaluator's results need to be with 1 point of each other
- where a difference exists, the two evaluators held discussions to achieve an agreed position[5]
- each evaluator makes changes to any of the other assessments based on that discussion
- this process is repeated until the entire batch of texts is processed

Each evaluator entered their assessments on an Excel spreadsheet (one spreadsheet for each text). The use of the instrument, as implemented in Excel, has been found efficacious by the assessors, allowing them to assess each text (of approximately one minute duration) within 20 minutes. Using an Excel macro, the results were subsequently aggregated onto a single

[4] These may overlap; so at worst case there are six texts chosen, at best three.

[5] If a position could not be arrived at, the process was to refer the matter to the research team, but in practice, the research team did not need to become involved at this stage of the process.

Volume 13, Number 3: December 2012
ISSN: 1535-0975

spreadsheet for subsequent analysis.

An aggregated score

With the overall objective of identifying 'high performing' classes, an aggregated score was computed to allow easy (but admittedly broad) comparison between classes. This aggregate score was calculated from that of the nine categories of meaning: those related to 'creating the world' (table 2), those related to 'showing a world' (table 3) and 'text structure'. That is, the scores for the non-essential component of 'other media', the non-semiotic 'technical innovation' and the more broadly interpreted 'multimodality' were omitted.

Inter-rater reliability

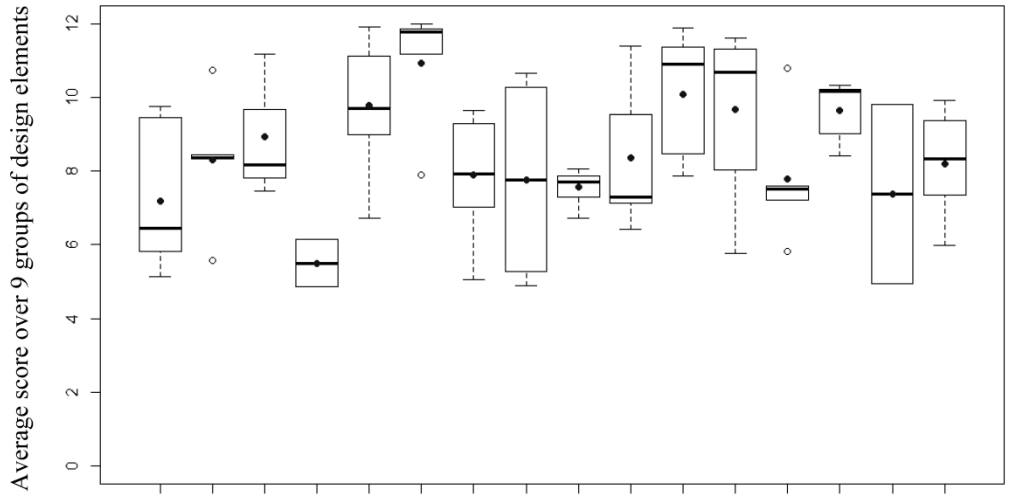
Several aspects of our work served to establish inter-rater reliability for the evaluators: the provision of the detailed evaluation guide, the training process for evaluators and the careful moderation process. Prior to proceeding with any further analysis, it was important to establish whether the results from one evaluator on the aggregated score could be distinguished from that of the other. In particular, using the statistical package *R* (R Development Core Team, 2012) the Mann-Whitney *U* test (Cohen & Lea, 2004, pp. 199-255) was applied to test the hypothesis that the sample from each evaluator has the same mean rank. The results for each evaluator for the aggregated score for the design elements has $p > 0.05$, thus the null hypothesis that the samples have the same mean ranks is not rejected, and so the results submitted by one evaluator cannot be distinguished from that of the other. This result enhances our confidence in inter-rater reliability, and further discussion and analysis has proceeded on this assumption.

Volume 13, Number 3: December 2012
ISSN: 1535-0975

Identification of high performing classrooms

Again using *R*, a box-and-whisker plot (Cohen & Lea, 2004, pp. 24-25) was used to present and compare the rating of aggregate scores for each class, as shown in Figure 2. On this figure, the median value is shown by the solid line and the solid dot shows the mean, the extent of the box describes the interquartile range. With a small sample from each class, the whiskers (which are designed to give a fair indication of the range of scores) and the outliers (shown by an open dot) add little to the meaning, except to indicate that the spread of results is fairly wide in some cases (A1, B1, B4, B6, D1, E3, E4). Visual inspection shows that performance of the texts selected from classes B1, B2, C1, D1 and E2 is quite striking.

Not only do the scores average in the 'high' range (i.e. greater than 9), but these are the only cases where the mean is distinctly less than the median. This indicates a left skewness, or more descriptively, there is a 'hump' of data at the top end. So whilst there were some texts in most other classes which performed quite well, the 'hump' is towards the lower end and would be so, even if the remaining 20-or-so texts from those classes were considered, as the best of them have been taken into account. We thus regard this set of five classrooms as 'high performing' because (a) the average scores are the highest of any (b) they are the only ones presenting as left skewed and a very high proportion of the best texts in this sample are found in these classes.



School:	A	A	A	A	B	B	B	B	B	B	C	D	E	E	E	E
Teacher:	1	2	3	4	1	2	3	4	5	6	1	1	1	2	3	5
Year Level:	5/6	5/6	5/6	5/6	6	6	5	5	5	6	6	5/6	5	5	6	6
No of texts:	6	5	3	2	7	7	8	4	6	5	5	8	5	5	2	5

Figure 2: Box plot comparing the performance of all classes

The boxplot shows the median by a solid line, the box indicates the interquartile range, and the mean is shown as a solid dot.

One needs to be careful in making assertions about what this might mean because the sample collected from each class is much too small to make any assertion about how the other texts from each of these classes may have performed. They could just as reasonably be evenly distributed along the whole continuum as clustered towards the top or the bottom of the scale. In the worst case, the distribution from each class would be bipolar, with a small sample of 'high quality' texts at the top and the remainder at the very low end; at best, they remainder accentuate the skewness already identified. Either way, it cannot be denied that the best texts in the sample will be found in those classes. By implication, then, we assert that there is something unique occurring in these classrooms. With a wider objective of developing an effective pedagogy of multimodal authoring, there is something to be taken notice of in classrooms such as these which can be investigated through different research methods (such as case study).

Volume 13, Number 3: December 2012
ISSN: 1535-0975

Discussion

Research into both multimodal pedagogy and the evaluation of students' multimodal texts is dominated by in-depth understandings of a small number of cases (Chandler-Olcott and Mahar 2003; Hull and Nelson 2005; Luce-Kapler 2007; Macken-Horarik and Adoniou 2008; Tan and Guo 2009; Mills 2010). The lack of ability to make broad assertions across a range of situations is a weakness. Whilst far from large-scale research, we have nevertheless been able to collect data from 17 classes comprising nearly 350 students, subject it to an evaluation method, and use that to identify specific classes to which greater attention might be paid. All these aspects are eminently scaleable to larger research ventures. Moreover, it allows for a dispassionate identification of classes which might be attended to in more detail. For instance, other parallel work (Thomas, 2011a, 2011b) would suggest that classes B2 and C1 would shed important light on multimodal authoring, while this data suggests that further examination of classes B1, D1 and E2 may be similarly illuminating. A close-up investigation of these classes is beyond the scope of this paper, but what we have sought to demonstrate is a proof-of-concept which can productively inform further work where the selection of classes is important.

There is an important caveat in relation to this group of 'high performing' classes, which arises from our observation of classes. It is this: it would be rash to think that the identification of a class as 'high performing' is solely related to depth or breadth of knowledge about meaning-making with multimodal texts. We have noted a range of inter-related factors which impact on that outcome: technical concerns which the school did not have any control of (i.e. bugs or consistencies within the software); technical concerns which the school might have been able to control, but for access to adequate technical expertise or resources (i.e. software installation or network performance issues, adequate numbers of

Volume 13, Number 3: December 2012
ISSN: 1535-0975

working computers); and classroom time-management issues, when the units of work took a much longer time than was estimated by the teachers. It is clear that something 'special' happened in some cases, but it is not possible to readily separate out circumstances which were strongly influenced by 'practical' concerns from those which arise from a high level of knowledge of the meaning-making strategies of multimodal authoring. Thus, a close-up investigation of 'high performing' classes may well tell a wider story of schooling, infrastructure, technology, support, collegiality and pedagogy.

Based on the success of the evaluation instrument to effectively discriminate between classes, its application to related endeavours is apparent. For instance to compare the performance on each category of meaning, and to consider whether student performance on each category of meaning is correlated with, or independent from, each other category. Such investigations could helpfully further inform a pedagogy of multimodal authoring.

Conclusion

The work described here sought to develop of an instrument that would facilitate the assessment of large numbers of multimodal texts produced by middle years students. This has been described in detail. At its core, a fairly simple idea has been implemented: identify a text as performing 'high', 'medium' or 'low' on 12 semiotically-derived criteria. To allow for ready comparison between texts, a range of numerical assignments, averages, aggregations and roundings have been applied. Through the trialling the instrument and all its computations in a meaningful investigation we have demonstrated a proof-of-concept, that such machinations do not reduce, summarise or dilute the data to a point that is ultimately unhelpful. That is demonstrably not the case. Through this approach, inter-rater reliability has been sustained, and the summary data can meaningfully distinguish between classes on the basis of the quality of the texts produced by the students.

Volume 13, Number 3: December 2012
ISSN: 1535-0975

Whilst we are encouraged by the success of the evaluation instrument, it is not a complete or final work in itself. Further development at both the theoretical and practical level are inevitable. Perhaps the most challenging conceptual issue is to better address collaboration between the modes, and perhaps the largest practical issue is the extent to which a 'distraction' rating is in fact an important inclusion. We have described a range of applications of this evaluation instrument, and present it as a work-in-progress which will contribute to understanding student construction of multimodal text and thus developments for effective multimodal pedagogy. The availability of instruments such as this will play an important role in the up-scaling of research ventures in this field.

Appendix 2: Evaluator Guidelines

Category of Meaning	Design element	Guidelines for considering the design element and the applicable conventions
Setting and location	Choice of location	Does the choice of location: Form a good basis for presenting the facts? Form a good basis for the mood and atmosphere desired? Form a good basis for the time/season/era desired? Permit appropriate choices with respect to a 3D capacity, and what might be desired in terms of the movement of participants, props, set dressing or the camera?
	Mood and atmosphere	Consider lighting, colour, fog and other special effects for mood and atmosphere. Consider the contribution of set dressing towards the mood and atmosphere.
	Temporal location (time of day, season, era, etc)	Some ways to achieve this might be: lighting (including intensity, direction and colour), the palette of colours used, the use of props or costumes, the use of set dressing and 'extras'. For instance, the presence of dinosaurs (an 'extra') would indicate an era as would a particular style of dress; events in a living room around an open fire may indicate evening. Considering the temporal location might not always be relevant because the choices situate the text somewhat nondescriptly, and this proves to be unimportant in relation to the meanings being conveyed (e.g. some retellings of nursery rhymes).
	Material location (environment, objects, inhabitants, etc)	In other words, the 'set dressing' and 'extras' (people) brought into the location for purpose of establishing the setting and location. Their purpose in helping establish location and mood are addressed above; here we are concerned with whether they "make sense". This includes such issues as whether explosions are used without apparent purpose or finding African animals in an Australian scene. (These are not the "things" which the story is "about" - see 'participant', below).
	Material processes (animation of environment, objects, inhabitants, etc)	This draws our attention to animations which should apply to set dressing and extras. If we want to convey the idea that "in the background, there was a koala is climbing the tree", we would animate the koala to actually climb the tree.

	Narration &/or dialogue	To include both the “sequence of words” and the performance of these. One approach to this is for a narrator to directly introduce the location of the story in time and in space, for instance “Once upon a time, but not very long ago, deep in the Australian bush lived two possums”. In dialogue, participants may make reference to location in time and space, which may be difficult to do visually, for instance “It’s been ten years since the war finished and we’re still living in this bomb-damaged hovel in worst suburb in east London”. A broad context could be provided verbally and then use the visuals to construct the immediate local context or juxtaposition.
	Sound effects	Sound effects may be used to convey information about temporal location (time/season/era), material location (the "things" in the location may be heard as well as seen, or heard instead of being seen) and physical movement (non-specific discussion between two people in the background may obviate the need to make them look as though they are talking).
	Background music	Music may be used to convey information about temporal location (time/season/era) and setting. For example, asian-style music helps establish the setting and medieval-style music would help establish an era.
Category of Meaning	**Design element**	**Guidelines for considering the design element and the applicable conventions**
Participant Selection and Construction	Material composition of a participant: Base choice of participant	The author/creator must choose the base participant, and whilst the subsequent change to colour, swatch, size, animation, etc can significantly impact on the basic object chosen, it still needs to be broadly suitable for the purpose for which it is appropriated. For instance, the use of a rhinoceros in an Australian scene is almost certainly inappropriate, regardless of how it is modified.
	Material composition of a participant: Color of participant, clothes, etc	The author/creator must choose whether the colourings/ swatches of the base participant are those which tell the story effectively. If not, then modification to the colour of the skin, hair or clothes should be made. For instance, a base participant may be a dancer, but because of positioning and colour of clothes may be made to look like a swimmer. A plain brown bird can be made to look like a magpie with appropriate swatching. A participant may be made salient by wearing a red shirt, or implicitly grouped with others as a family because they all have clothes swatched the same colour.

	Material composition of a participant: Relative size and shape	The author/creator must choose whether the size and shape of the base participant are those which tell the story effectively. If not, then modification to the shape and size of the participant should be made. For instance, an egg may be the size of a hen's egg, relative to other participants, but should be made larger because an ostrich egg is under consideration, or needs to be more prominent. A story of a 'roly-poly policeman' needs to show someone who is not as thin as a rake.
	Material composition of a participant: Animation	This is the use of internal animation to show more about "what the participant is like" or "how the participant interacts with others" than would otherwise be apparent through the base participant alone. It might also be keypointed animation in the sense of conveying that someone has a limp, a strong throwing arm or an elegant dancer.
	Narration &/or dialogue	Voice performance (in particular, words of dialogue 'from' the participant, 'to' the participant or narration 'about' the participant) may be used to show more about "what the participant is like" or "how the participant interacts with others" than would otherwise be apparent through the visual presence of the participant alone. This includes both the "sequence of words" and the performance of these. For instance, a gruff voice or being spoken about in sarcastic tones. Very occasionally the participant may not be visually present at all, but known to the viewer only through sound effects and voice performance.
	Sound effects assigned to a character	Sound effects may be used to show more about "what the participant is like" or "how the participant interacts with others" than would otherwise be apparent through the visual presence of the participant alone. For instance, heavy footsteps or burping. Very occasionally a participant may not be visually present at all, but known to us only through sound effects, possibly by giving a participant a 'signature sound'.
	Background music assigned to a character	The possibilities are limited. An example would be giving each participant a signature theme, such as in Prokofiev's 'Peter and the wolf'. In that way, "what the participant is like" or "how the participant interacts with others" can be conveyed more richly than would otherwise be apparent through the visual presence alone.

Category of Meaning	**Design element**	**Guidelines for considering the design element and the applicable conventions**

Category of Meaning	Design element	Guidelines for considering the design element and the applicable conventions
Arrangement and Interaction of participants	Physical placement	This refers to how the participants/objects are placed "physically". For instance, do participants who are supposed to be standing on the ground look like their feet are actually on the ground? This category focuses our attention on whether there are "problems" of this type, or whether it has been handled relatively well. Whilst in theory participants are always located physically, sometimes this is not a relevant consideration, for instance in a series of close-up shots it may not matter if physical positioning has flaws or not. Also consider whether necessary “re-adjustment” of physical placement has been managed well. For instance, a base participant may be a dancer, but because of positioning and colour of clothes may be made to look like a swimmer.
	Physical arrangement	This refers to the "actual" layout. It is more likely that the viewer will see what the camera wants us to see, rather than what is really there, which is addressed in the 'apparent visual layout' (below). But just in case there is a shot from which the viewer can correctly determine the "actual" layout, this is where it should be judged. Have there been, for instance, deliberate use of centre/periphery or ideal/real, new/given? Particular gazes, animations or vectors to connect the people/objects in the scene?
	Material processes of participants (interaction)	In this, we are concerned with: how the participants move "actually" (i.e. in a ball-room scene, it would be conveyed that who participants are dancing together because they are seen to move towards each other and are animated with dancing gestured), varying social relations between participants, or participants and a viewer, varying salience - changing which participant is highlighted by change of positioning from one moment to the next. There may be times when interaction is not a relevant consideration, for instance if there is only one participant. (Note that the physical movement of props, set dressing etc is separate, as above).
	Narration &/or dialogue	This includes both the “sequence of words” and the performance of these. Same kind of issue as above - more likely to show via image construction.
	Sound effects	The presence of sound effects, dialogue or narrative may identify other participants who may or may not be clearly identified otherwise, and whose action in relation to others may not otherwise be identified. For instance, a presumed off-screen golfer yelling "fore!" prior to a golf ball knocking a participant to the ground.

Category of Meaning	Design element	Guidelines for considering the design element and the applicable conventions
Sequencing of information	Sequencing of scenes *(visual mode only)*	The scenes are actually in the correct (narrated/logical) order The (implied) time sequence is sensible. For instance, if a closing scene is depicted as late afternoon, it would be expected that scenes apparently occurring earlier in the same day would be shown in daylight rather than dark.
	Ordering of shots *(visual mode only; must of have least one scene with two shots)*	The normal state of affairs is for one shot to show one action/transaction/event, and the next shot to show the next one ('temporal sequentiality'), with other possibilities such as 'temporal simultaneity', 'flashback')
	Pacing	A very literal understanding of pacing is meant here. For instance, the text might identify a participant through a close-up shot, but the camera might 'wait' on that participant for insufficient time for the viewer to make the connection that a participant has been identified. Alternatively, the text might move too slowly – the author has not seemingly made adjustments to the duration of the scenes/shots to adequately reflect the mood or atmosphere otherwise intended.
	Narration &/or dialogue	This includes both the "sequence of words" and the performance of these. Narrative or dialogue could be used to provide information of events intervening between one scene and another so that the verbal information was what maintained continuity.
	Sound effects	Examples are a clock striking 'on the hour' which to show an appropriate progression of hours, or the sound of a train "off stage" signalling that either someone had departed or was about to arrive.

Category of Meaning	Design element	Guidelines for considering the design element and the applicable conventions
Viewer stance	Point-of-view	The point of view conventions are: • The 'default' position is that of external observer with no connection to the action. The viewer is positioned to see what is happening. • Directly as the viewer, for example through the position of the camera which places the viewer in amongst the action. • Along with character where the viewer is positioned with an over shoulder view, or close behind or beside the character seeing part of the body, but also seeing what the character is seeing. • As a character first person point of view, for example indicated by hands/feet/shadow in view to indicate the camera is 'someone'. First person point of view can also be established through a shot- reverse shot sequence which infers the viewer is seeing the action as a character. Shot one shows the action from a characters point of view, shot 2 shows the characters reaction. This can also work in reverse.

Category of Meaning	Design element	Guidelines for considering the design element and the applicable conventions
The selection of visual information (framing)	Camera distance	Camera distance is used to portray various degrees of social distance through wide shot, mid shot, close up or extreme long shot.
	Hide/Reveal	Hiding/revealing is concerned with: • disclosing certain information which is actually 'in the scene' yet keep other information hidden (ie hide/reveal) • viewing a group in such as way as to infer that they are grouped together (whether they are nor not). For instance, if 3 people are sitting around a table, it can be inferred that two of them are 'close' and the other excluded.
Angles through which visual information is seen	Vertical camera angles	Vertical camera angles are used to portray various degrees of social power, especially between the viewer and the participant. For instance, the use of eye-level, low angle, high angle, bird's eye view. These also imply some sense of point-of-view.
	Horizontal camera angles	Horizontal camera angles are used to portray various degrees of involvement, especially that between the viewer and the participants. For instance, the use of frontal angle, oblique angle, back view or bird's eye view.
Movement of the viewers with respect to that which is viewed	Camera movement	The codes here are zoom, tilt, plan tracking and fly. These can be used to use to dynamicize social distance, social power and involvement Also consider: how the speed of camera movement can influence meaning whether a filmic cut would actually be more appropriate
Multimodality	Collaboration between visual/verbal/aural	In other words, it shows rather than tells. It is essential to refer to "whether sufficient information is communicated to tell the story"; there must be sufficient information in various modes actually communicated before a judgement can be made with respect to this. An example of a collaboration: a 'composition conjunction' would show disparate activities, but because of their temporal sequence, we understand them as similar or related. A consistent voice over (or continuity of music) could confirm this.
Other media	On-screen text	One must be reasonably sure that the use of an audio track would not be a better option, but since on-screen text is available this is the place to make judgements about it.
	Titles, credits, etc	Consider the choice of font style, colour/s, size and position on the screen for the titles and credits, work with or support the story genre, as well as enhance the aesthetics or look and feel of the text
Technical innovation	Significant technical innovation	Some examples may be: Sophisticated grouping/swatching to produce space ships, creating rain in a scene, creating a shadow

Category of Meaning	Design element	Guidelines for considering the design element and the applicable conventions
Structure of the text	Is there sufficient information to tell a story	The viewer is given enough relevant information about what is happening in this story for it to make sense.
	Orientation	Orientation is the introduction of the characters, location in time and space and identification of any key elements of context significant for the story.
	Complication	Complication is the problem or issue that arises that disrupts the routine that normally prevails.
	Evaluation	Evaluation is the participant's reflection or judgments about what is happening (to them).
	Resolution	Resolution is the solving of the problem, resolution of the issue, return of the situation to normalcy or a new equilibrium.

References

Anstey, M. and Bull, G. (2006). *Teaching and Learning Multiliteracies: Changing Times, Changing Literacies.* Newark, JN: International Reading Association.

Australia, National Curriculum Board. (2009). The Shape of the Australian Curriculum: English. Canberra, Commonwealth of Australia.

Baxter, D. and Simpson, A. (2008). Secondary School English, Literacy Examinations and Attitude: Utilising Appraisal to interpret values in image/text relations. In L. Unsworth (Ed.), *New Literacies and the English Curriculum: Multimodal Perspectives* (pp. 309-327). London: Continuum.

Burn, A. and Parker, D. (2003). Tiger's Big Plan: Multimodality and the Moving Image. In C. Jewitt & G. Kress (Eds.), *Multimodal Literacy* (pp. 56-72). New York: Peter Lang.

Burn, A. and Durran, J. (2006). Digital Anatomies: Analysis as Production in Media Education. In D. Buckingham and R. Willett (Eds.), *Digital Generations: Children, Young People and New Media* (pp. 273-294). Malwah, New Jersey: Erlbaum.

Burn, A. and Leach, J. (2004). ICT and moving image literacy in English. In R. Andrews (Ed.), *The Impact of ICT on Literacy Education* (pp. 151-179). London and New York: Routledge Falmer.

Chandler, P. (2010). Not always as it first seems: thoughts on reading a 3D multimodal text. *Literacy Learning: the Middle Years, 18*(1), 11-18.

Chandler, P., O'Brien, A., & Unsworth, L. (2010). Towards a 3D digital multimodal curriculum for the upper primary school, *Australian Educational Computing, 25*(1), 34-40.

Chandler-Olcott, K. and D. Mahar (2003). "Tech-saviness" meets multiliteracies: Exploring adolescent girls technology-related literacy practices, *Reading Research Quarterly 38*(3),

356-385.

Cohen, B. H., & Lea, R. B. (2004). *Essentials of statistics for the social and behavioral sciences*. Hoboken, NJ: Wiley.

Collerson, J. (1997). *Grammar in teaching.* Newtown, NSW: Primary English Teaching Association.

Cope, B., and Kalantzis, M. (2009). A grammar of multimodality. *The International Journal of Learning, 16* (4), 361-426.

England, Qualifications and Curriculum Authority. (1999). The National Curriculum for England; English. Department of Education and Employment. London.

Halliday, M. A. K. (1978). *Language as a social semiotic: The social interpretation of language and meaning*. London: Edward Arnold.

Halliday, M. A. K. and Matthiessen, C. (2004). An *introduction to functional grammar*. London: Arnold.

Harris, M. (1979). *Cultural materialism: the struggle for a science of culture*. New York, NY: Random House.

Hull, G. and Nelson, M. (2005). Locating the semiotic power of multimodality. *Written Communication 22*(2), 1-38.

Kimber, K. and Wyatt-Smith, C. (2008). Assessing Digital Literacies: Can Assessment Ever Be the Same? In L. Unsworth (Ed.), *New Literacies and the English Curriculum: Multimodal Perspectives* (pp. 328-352). London/New York, Continuum.

Kress, G. & van Leeuwen, T. (2001/2006). *Reading Images: A grammar of visual design.* London, Routledge.

Luce-Kapler, R. (2007). Radical Change and Wikis: Teaching new literacies, *Journal of*

Adolescent and Adult Literacy, 51(3), 214-223.

Macken-Horarik, M. (2008). Multiliteracies and 'basic skills' accountability. In L. Unsworth (Ed.), *New Literacies and the English Curriculum: Multimodal Perspectives* (pp. 367-382). London/New York, Continuum.

Macken-Horarik, M. & Adoniou, M. (2008). Genre and Register in Multiliteracies. In B. Spolsky and F. M. Hult (Eds.), *The Handbook of Educational* Linguistics (pp. 367-382), Malden, MA: Blackwell Publishing.

Maggs, P. (2008). Kahootz 3.0: Developing software for the classroom. *Teacher, March*, 28-31.

Mills, K. (2010). Filming in Progress: New spaces for multimodal designing, *Linguistics and Education 20*(1), 14-28.

Mills, K. (2011). iPed : pedagogy for digital text production, *The Reading Teacher, 65*(1), 80 - 91.

Mills, K. (2011). Now I know their secrets: Kinekonic texts in the literacy classroom, *Australian Journal of Language and Literacy, 34*(1), 24-37.

R Development Core Team. (2012). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing Vienna Austria. Retrieved from http://www.r-project.org

Singapore, Ministry of Education. (2010). English Language Syllabus 2010 Primary and Secondary (Express/Normal[Academic]).

Tan, L. & Guo, L. (2009). From Print to Critical Multimedia Literacy: One Teacher's Foray Into New Literacies Practices, *Journal of Adolescent and Adult Literacy, 53*(4), 315-324.

Thomas, A. (2008). Machinima: Composing 3D multimedia narratives. In L. Unsworth (Ed.), *New Literacies and the English Curriculum: Multimodal Perspectives* (pp. 167-185).

London/New York, Continuum.

Thomas, A. (2011a). Children's Writing Goes 3D: A case study of one school's journey into multimodal authoring. *Learning, Media and Technology*. EJ ISSN 1743-9884

Thomas, A. (2011b). Developing a transformative digital literacies pedagogy. *Nordic Journal of Digital Literacy*, *6*(1-2), 89-101.

Unsworth, L. (2008). Multiliteracies, e-Literature and English Teaching, *Language and Education 22*(1), 62-75.

Unsworth, L. & Chan, E. (2008). Assessing integrative reading of images and text in group reading comprehension tests, *Curriculum Perspectives, 28*(3), 71-76.

Unsworth, L. & Chan, E. (2009). Bridging multimodal literacies and national assessment programs in literacy, *Australian Journal of Language and Literacy*, *32*(3), 245-257.