Joining the Dots: The Literacies of Multimodal Longform Journalism

Andrew Duffy

Routledge
Taylor & Francis Group

RESEARCH ARTICLE

Joining the Dots: The Literacies of Multimodal Longform Journalism

Andrew Duffy

Wee Kim Wee School of Communication and Information, Nanyang Technological University, Singapore, Singapore

ABSTRACT

Each emerging form of communication demands a new or at least adapted form of literacy. What remains constant is the user's ability to critically analyse messages in whichever form they take. Multimodal longform journalism requires compound literacies to access, read, process, and make meaning. This article examines how meaning is made when readers navigate a complex, multimodal longform story which requires them to toggle between text, image, video, graphic and hyperlink. It describes the novel approach of autoethnographic textual analysis and delivers a 'report on experience' of immersion in five complex multimodal longform stories, followed by discussion of the literacies required to create and consume multimodal longform, and the implications for scholarship.

Literacy has many variations (Wu and Wang 2011) and each new form of communication demands a new or at least adapted variant: blogging requires new skills of interpretation, as do photo and video sharing sites, for example, and even customer review sites (Pilgrim and Bledsoe 2013). What remains constant is the user's ability to critically analyse messages, whether text, video, image, audio, or on multidimensional, multidirectional channels.

The emergence of new channels for journalism and its delivery through image and sound alongside text (Thoman and Jolls 2005) has led to renewed interest both in its literacy (Silverblatt et al. 2014) and in the way that its literacy is assessed since "each narrative medium requires a specific analytical approach to narrative structures and levels" (Onega and García Landa 2014, 2). The technologically driven phenomenon of multimodal longform digital journalism epitomised by the *New York Times*' iconic 'Snow Fall: The Avalanche at Tunnel Creek' (Branch 2012) requires compound literacies to access, consume, process and make meaning. This study offers a novel analytical method to investigate these compound literacies.

Since 2013, multimodal longform stories have become a means for newsrooms to showcase their abilities. 'Snow Fall' "established a model various media outlets have

---

since mimicked to cultivate a similar sense of prestige" (Dowling and Vogan 2015, 209).

This article investigates the descendants of 'Snow Fall' to consider how meaning is made when audiences navigate a complex, multimodal longform story which requires them to toggle between text, image, video, graphic, and hyperlink. It asks a first research question: given the complexity of these stories, what is the process by which a reader engages with them? This question becomes more pressing as complexity of these stories increases, bringing ever more demands on the reader. The next level might be immersive journalism which uses virtual reality and 360-degree video to create even deeper engagement (Sánchez Laws 2020) and empathy (Hassan 2020), indicating that the need for even more advanced media literacy when engaging with complex journalism will continue to evolve.

At its most basic level, multimodal journalism involves three types of grammar: textual, graphic and interactive (Palau-Sampio and Sánchez-Garcia 2020) which in turn demand three literacies. A further challenge is to co-ordinate them. While literacy scholarship has identified critical thinking about media forms, less well studied is the way these interact with each other to create an effect which in itself demands critical thinking, and thus an extension of theorising into literacy. This article investigates meaning-making in multimodal longform digital journalism, and asks a second research question: how does multimodality and interaction between modes impact on the way the story is interpreted by a reader? It looks at the literatures on literacies, multimodality, and longform narrative journalism. It describes the novel approach of autoethnographic textual analysis of five longform stories, followed by discussion of the literacies demanded. Finally, it poses a third research question: when a reader joins the dots of a multimodal narrative, what picture appears of the literacy involved when engaging with such a story?

The Idea of Literacy

Few disagree that literacy is essential; more dispute what it *is* (Wallis and Buckingham 2019; Pilgrim and Martinez 2013). As a starting point, this article follows Coiro and Dobler (2007, 217) who saw it as "an active, constructive, meaning-making process … [where] readers actively construct meaning as they interact with text … [and] use a range of strategic cognitive processes to select, organize, connect, and evaluate what they read." Beyond this, literacy is concerned with relationships between three processes: the representation of culture, ethics and knowledge using symbols; the spread of skills to interpret media messages; and the institutional management of power which comes to those who are literate or have the ability to access and use knowledge (Livingstone 2004). At the heart of literacy, therefore, is critical thinking: to be literate in a domain is to be able to evaluate, interpret and critique what one is witnessing, reading or experiencing in order to make meaning from it and ultimately to guide decision making (see Potter 2004; Maksl et al. 2017; Craft, Ashley, and Maksl 2016). This article examines how multimodal longform is evaluated and interpreted and how meaning is made through interaction with and between modes making up the story.

News and Media Literacy

Media literacy involves knowledge about the functioning of mass media in society. Its purpose is to mediate "the complex relationship that people have with media, popular culture and technologies of communication" (Hobbs 2007). It incorporates news literacy which can be defined as "knowledge and skills related to news production, distribution, and consumption that more closely aligns with public understandings of the term, creating both constancy in its conceptualization and flexibility in its application" (Vraga et al. 2021, 13). It can motivate citizens to be better informed, reflective, and engaged in civic matters which can be a boon for democracies (Mihailidis and Thevenin 2013); or the opposite, as learning about media can make users disengaged, cynical and apathetic (Ashley, Maksl, and Craft 2017).

New technologies of communication demand new literacies and "require different ways of conceiving and communicating meaning presented in multiple media and modality forms as a part of literacy … new web-based tools emerge on the Internet constantly and require specific, new skills" (Pilgrim and Martinez 2013, 63). This kind of digital literacy is complex; it involves "cognitive, motor, sociological, and emotional skills" (Eshet-Alkalai 2004, 93). It may be in short supply, and only 17% of adults in the US are 'digitally ready' and are confident and skilled to acquire new information online (Horrigan 2019).

Multimodal Longform Is Different

This raises the question of how skilled people are in engaging with multimodal longform. A multimodal text is a complex entity that tends to exist in the digital realm and uses varied cultural and semiotic elements to convey complex and sophisticated ideas (Serafini 2014). It "combines written language, photography, short videos, maps and other graphical elements, and joins them together into a seamless narrative using subtle transitions [and] seeks to captivate its audience by combining text, photographs, looping videos, dynamic maps and data visualizations into a unified whole" (Hiippala 2017, 420).

Scholarship started with Knox (2007) analysing 'newsbite' stories on the landing pages of news websites with a headline, brief text and an image, created for speedy consumption. Since then, other studies have looked at how cinematic and literary it is (Jacobson, Marino, and Gutsche, 2016), the role of multimedia at its heart (Hiippala 2017), and user engagement (Dowling and Vogan 2015). While founded on news reporting principles, it goes far beyond the norms of news. It is characterised by increased experimentation beyond the restrictions of text using hypertextuality and driven by creativity partly in the pursuit of innovation, partly in the pursuit of legitimacy and credibility, and partly in the pursuit of a dwindling revenue sources to keep journalism running (Vázquez-Herrero, López-García, and Irigaray 2020).

To create multimodal stories, journalists must be multiskilled; to consume them, readers must multitask (Menke 2019). While journalists must make the story immersive, the audience must make time to immerse themselves in it. Both place off-deadline, multimodal longform at the far end of a spectrum from bite-sized breaking news which demands immediacy and simplicity. It also requires broadband Internet access

to access its complex, data-heavy content at a reasonable speed, raising the issue of accessibility: multimodal longform is not for everyone.

According to Dunham (2020, 4) this form of journalism "has the same rules for accuracy, fairness and ethics" as traditional news reporting. This article questions such characterisation, however. Instead, it suggests that ethics, accuracy, fairness, neutrality and above all objectivity may be sacrificed in the pursuit of an engaging narrative. One goal specific to multimodal longform may be transportation which occurs when the audience "lose track of time, fail to observe events going on around them, and feel they are completely immersed in the world of the narrative" (Green, Brock, and Kaufman 2004, 247). Clearly, a consideration of news literacy alone is insufficient for analysis of this sophisticated form.

Literacies Converge

The idea of multiple literacies is used to describe the way people work with diverse forms of communication such as books and newspapers, film and music, speech and gesture, each of which requires particular skills or understanding to interpret successfully (New London Group 1996). Multimodal longform requires readers to toggle back and forth between modes: text, image, video, graphic, audio, hyperlinks and interactive affordances. Each of these demands a form of literacy. Yet there is an additional literacy required to *assemble* the disparate modes into a coherent narrative; to be critically aware of how the modes were set up to interact with each other in order to create a certain effect; and to be critically reflexive of one's own process in navigating through the multimodality and the conclusions one reaches.

Digital multimodal longform therefore demands both multiple literacies *and* the ability to assemble literacies to make meaning. Multimodal literacy is "a process of generating meaning in transaction with multimodal texts, including written language, visual images, and design features, from a variety of perspectives to meet the requirements of particular social contexts" (Serafini 2015, 413). The elements or modes of a story have been selected, created and edited to work together; it is therefore reasonable to analyse them as such, to 'join the dots'. On the one hand, "components of the story are crafted to complement one another" (McAdams 2015, 188) but on the other they can equally distract attention or fragment the experience (Giles and Hitch 2017).

> Multimodal forms of information include visual and audio modes of communication presented through print, photos, videos, or graphs. These various modes of communication affect the way readers approach text. For example, graphic novels … require visual literacy skills to comprehend both the text and the illustrations used by the author to represent meaning" (Pilgrim and Martinez 2013, 62; Kress 2010).

This misses out that the audience must also make sense of the interaction between modes – images and words are not interpreted separately but in conjunction with each other and are constructed that way to create a desired effect. As a result this article also studies the *interaction* among disparate modes of multimodal longform as a crucial aspect of media literacy.

Interaction between modes suggests a narratological perspective (Brooks 1984; Darby 2001). Beyond the simple linear flow of a narrative – what event is represented

as occurring first, second, third and so on – such a perspective considers the ways a narrative is a compound of elements, looking at the structure of the representation of events as well as the events represented (Onega and García Landa 2014; Guillemette and Lévesque 2019). A narratological perspective therefore allows for analysis of how the way in which a story is structured impacts the way it guides an audience to perceive both the story itself and the wider world it refers to, and the meaning created from the narrative as it is presented.

An Analytical Model of News Literacy

Such a complex undertaking benefits from a guiding framework. Potter (2004) proposed a cognitive model of news literacy with four connected components. The first involves an understanding of what the media industry does and how it achieves its goals through choice of content and presentation. Second is a personal locus, which can override heuristic processing of news media and drive motivation to engage with the narrative. The third is the competencies and skills demanded to process the media in its varied forms. And the fourth component is how media is processed in order to manufacture meaning or match it to existing opinions or schema. Potter (2004) argues the benefit of his model is that understanding messages and the motives of those who produce media allows us to choose those with a positive influence and avoid those whose influence is less benign.

Craft, Ashley, and Maksl (2016) built on this framework, observing that it suggests a level of cognition which may not be present in all consumption of news media. Many people take a heuristic rather than cognitive route (Tversky and Kahneman 1974). Yet, given that multimodal literacy involves developing cognitive skills, it is a valid structure for this analysis. This study therefore considers four research areas corresponding to Potter's components and offering a structure for analysis of how multimodal longform is read.

Method: Autoethnographic Textual Analysis

The disparate elements of multimodal longform are "inherently structured in a way that provides their user with a range of choices, and the selections made within these choices may be combined into more complex structures [and] require the viewer to resolve discourse relations that are primarily spatial, that is, determining which pieces of content belong together" (Hiippala 2017, 422). Each person assembles the parts to make meaning and the outcome may vary between individuals. What is important for this study, as in the first research question, is the *process* of assembling elements or modes of a narrative to make meaning, rather than the resulting meaning itself. Further, as in the second research question, observing the process allows researchers to view the interaction between the elements or modes and how it impacts on the way the story is interpreted.

This study therefore proposes *autoethnographic textual analysis* to report on an experience of a reader consuming multimodal longform, determining how modes are assembled to make meaning by one individual, and what the meaning of that

meaning is. Autoethnography is a "theoretical, methodological, and (primarily) textual approach that seeks to experience, reflect on, and represent through evocation the relationship among self and culture, individual and collective experience" (Holman Jones 2007, n.p.). It "uses personal experience ('auto') to describe and interpret ('graphy') cultural texts, experiences, beliefs, and practices ('ethno')" (Adams, Ellis, and Holman Jones 2017, n.p.). The method has been used extensively to report from inside a community or shared experience; this article extends this use into reporting on a cultural experience.

Autoethnographic textual analysis is a novel method and novelty needs justification. This article argues it has three advantages. First, it is congruent with the topic under scrutiny: autoethnography has been used to investigate cultural issues as wideranging as eating disorders, creativity, gender norms, sport, adoption, swearing and racism. This article argues that storytelling is a cultural practice and that longform journalism with its predilection towards stories about particular cultural groups or issues emphasises this attribute. Literacies also include cultural considerations when making meaning in a text (Cervetti, Damico, and Pearson 2006).

Second, much multimodal longform pursues a convention of ethnographic realism which works from the starting point that subjective reporting can be a truthful representation of reality (Ward 2004). An alternative perspective is that journalists operating in the cultural phenomenology paradigm work from the starting point that the interplay between the external event and the internal world of the individual (journalist) is where factuality is constructed. As a result, ambiguity is hard-wired into this form (Eason 1984). Either way, this form of journalism occupies a place both in a culture it reports *on* and the culture it reports *for*; as a cultural artefact, it is therefore congruent with a cultural approach concerned with how its reality is constructed by an audience.

Third, the narrow focus inherent in autoethnographic textual analysis allows for in-depth analysis, unmediated by interpretation of a third party. This also replicates the individual's natural way of consuming multimodal longform: alone, and filtering it through their own cognition, their own heuristics, and their own schemata.

Additionally, this method invites reflection on the process of reading, making it suitable to investigate the first research question which concerned just such a process. It invites analysis of how meaning is made in the interaction of the different modes, making it a relevant approach to answer the second research question. And because the goal is analysis *of the process* to gain insights into literacies demanded by multimodal longform, it also tackles the third research question of what picture of literacy emerges when one joins the dots.

That is not to say the method has no downsides. Thomas (1993) identifies pitfalls in autoethnography generally, including seeing only what supports the intention of the study, becoming emotionally involved rather than scientifically dispassionate, and making claims that go beyond the data. This, however, suggests that a positivist ontology of neutral objectivity is still held up as the goal of study, which is to misunderstand the subjective, descriptive aim of autoethnography. Delamont (2009), meanwhile, has criticised autoethnography as intellectually lazy and critiqued it on the grounds that research should be analytical rather than experiential. Weir and Clarke (2018) refute this, arguing that the work is 'tough', and its function is to use the experiential as a departure point for analysis.

Nevertheless, there is a need to deal with validity, reliability, rigour and legitimacy of autoethnography with reflexive care (Sparkes 2002; Duncan 2004). That involves recognition of the self as the centrepiece of any autoethnographic textual analysis as much as the topic studied, while avoiding solipsism so that a study must return to the phenomenon experienced rather than only the experience of the experience (Wall 2008). Reflexivity cannot simply turn back on itself or it risks self-absorption; it must undergo a 360-degree turn to be directed towards the original phenomenon or issue (Ettema 2005; Siegle 1986). This article therefore attempts to be rigorous first in its report on the subjective experience of reading and subsequently in reflecting on that experience.

A Three-Round Process

Autoethnographic textual analysis is proposed as a three-round process. In round one, a researcher reads a text while audio-recording their thoughts about the process. This is transcribed into a *report on experience* describing how they engaged with the text and what they felt about doing so; and how they reacted to the activities of reading, scrolling, clicking and making meaning from the disparate narrative modes as they connected with each other.

For round two, this report on experience is subjected to thematic content analysis which is an established "research technique for making replicable and valid inferences from texts … to the contexts of their use" (Krippendorff 2011, 24). The second round also involves critical reflection on the significance of these themes for the researcher, identifying biases and assumptions on the grounds that autoethnography demands reflexivity, an awareness of the researcher's role and how it is influenced by the topic (Haynes 2012). It also concerns the capacity to see influences and to change actions, behaviour or even personal image as a result (Ahva 2013).

Round three of analysis builds on the first two: the researcher returns to the original text with the critical awareness of the first two rounds to observe which themes (from round two) based on the report on experience (round one) become more or less significant during a more cognitively aware reading, and describing them in detail. Ultimately, the goal of this research method is to suggest critically aware insights into a cultural phenomenon, in this case the literacies of multimodal longform.

Sample Selection

Examples were sourced by googling 'award-winning multimodal longform journalism'. The first web page returned was "5 great examples of innovative longform journalism online" a selection of an editorial analytics company (SmartOcto 2020). The stories were Riding the New Silk Road (*New York Times*), The Rise of the Deepfake (*Guardian*), China's Hidden Camps (BBC), Scientology's Clear Takeover (*Tampa Bay Times*), and Scaling Everest (*Washington Post*). They were selected for this study not as representatives of all forms of multimodal journalism but as a starting point to test the approach of autoethnographic textual analysis and to indicate issues of the literacies demanded by multimodal longform. Looking beyond the first website returned, other collections

of highly rated multimedia longform were also noted, such as the Online Journalism Awards' annual winners. Yet its number of stories awarded (54 winners and 67 finalists across 19 categories) was impractical for a single-author autoethnographic study, and so the five stories from SmartOcto were selected as adequate.

To give a brief description of each story:

- Riding the New Silk Road (New York Times 2013) uses graphic maps, photographs, videos and captions to show how Hewlett-Packard transports computer parts by rail from China across the steppes of Kazakhstan to Europe.
- The Rise of the Deep Fake and the Threat to Democracy (Guardian 2019) leads with a video of US TV host Jimmy Fallon in a sketch playing then-US president Donald Trump, followed by a video clip showing how it was transformed into a deep fake. One side of the screen carries text with hyperlinks to background information; on the other side video and audio clips, photographs and captions pop up as the reader scrolls down, illustrating the variety and impact of deep fake technology.
- The investigative story China's Hidden Camps (BBC 2018) leads with a satellite image of buildings in China's Xinjiang Province which it claims show detention camps (although China calls them vocational schools or re-education schools) for Uighur Muslims. Scrolling down reveals a series of images of people in the location and text explaining what the reporter found. Time-stamped maps change to show the development of desert into built-up land.
- Similarly investigative, Clear Takeover: How Scientology Doubled its Downtown Footprint in 3 Years (Tampa Bay Times 2019) uses Google Earth satellite photography superimposed with graphics highlighting changes over time, with photography, graphics and text showing the changes at street level and hyperlinks to related stories.
- Scaling Everest (Washington Post 2016) has an illustrated, monochrome graphic representation of Mount Everest. The reader scrolls upwards from 'sea level' to read and see what is present at each elevation, as well as practical, cultural or environmental information, accompanied by audio clips from people who have made the ascent.

Round One: Report on Experience

The first round of analysis involves the researcher reading the texts and reporting on their experience. In this regard, the first round is not dissimilar from watching a movie or reading a novel and telling a friend about it, albeit done with more cognitive critical awareness. Students may be asked to write book reports; newspaper critics report on cultural events; both are writing a form of report on experience. The difference here is that autoethnographic textual analysis takes this report through two further rounds and has an analytical end-goal in mind – here, describing the literacies required to engage with multimodal longform. The researcher read each article (8 to 20 min per article, $M = 15.2$ min) and audio-recorded their thoughts about the process of reading. Literacy is a combination of applied skills and critical thinking; it involves

evaluating and making meaning; and consideration of the symbolic representation of culture, ethics and knowledge (Livingstone 2004); these three concepts offer a framework for the presentation of the report on experience.

Applied Practical Skills

The basic literacy skills of critically responding to text, image, graphics, video and audio have been well documented and do not need restating here. The researcher did, however, note the practical skills needed to interact with the multimodal longform. Simply, scrolling requires the skill of maintaining eye contact with the end of one sentence as it moves up the page, and moving easily to the beginning of the next sentence. This movement involves a single leap using the 'page down' function, or blurring the text using the 'scroll down' function. A second form of interaction is the skill of leaving one mode to shift to a different one (e.g. text to audio, graphic to text) and then back again. Third, the researcher noticed that each mode found meaning in relation to another, until all the parts were assembled in the researcher's mind into a single, unified story.

Evaluating, Critical Thinking and Making Meaning

While processing each story, the researcher noted that they were also thinking about the reputation of the news organisation. Preconceptions about the news organisation's values – such as presumed political bias or neutrality – therefore influenced the meaning of each story. This raised the thought that most readers are likely to encounter multimodal longform on the organisation's proprietary platform; and readers are likely to be on that platform because they enjoy what they read there. Any cultural or political biases observed in a story may therefore coincide with the reader's own. Additionally, the researcher recognised that extra effort – time, money, personnel – had been put into each story; this was not the newsrooms' standard reporting approach.

Representing Culture, Ethics and Knowledge

The researcher also observed that each story used shared cultural reference points that linked what is new to what is known. These are described in more detail in round three. The researcher was aware that they were being managed in the way in which they fitted the information into their personal schema through the way the stories were framed. Nonetheless, it is possible that the researcher did not recognise some reference points that other readers would see and appreciate.

Round Two: From Reflection to Themes

Significantly, the researcher was aware of the pleasure in interacting with the stories in each of the three ways set out above: the rewards of engaging with each mode; of critical thinking; and of sharing cultural reference points; and of the extra effort

expended by the media organisation to create this effect, which they often referred to.

This level of cognitive recognition of the skills needed and their effect on the reader is a central principle of autoethnographic textual analysis, which makes cognitive, active and conscious an activity which is often heuristic, passive and unconscious. It involves critical thinking about the process of reading, clicking, and interacting with and combining modes; and critical thinking about the product itself, including the tactics that suggest an agenda and create a reaction. Its value lies in the abstraction of the individual's interpretation which may have more generalisable value, rather than the interpretation itself which may be specific to the individual and the context.

The audio-recorded and transcribed observations from the report on experience were therefore subjected to the second round of textual analysis. This analysis gave rise to abstractions of seven themes related to literacy: media showcasing its work; reader expectations; agendas in evidence; interaction with modes; reward and pleasure; connecting the elements; and interpretation. These allied to the four components in Potter's cognitive model of news literacy (2004), which is therefore used as a framework to describe the themes, presented in more detail next, in the third round of analysis.

Round Three: Detailed Description

Potter's first component is how much a reader understands what the media does and how it achieves its goals through choice of content and presentation. This was observed in three themes related to literacy identified in round two and are described in detail here:

Media Showcase

The researcher observed that the news organisations used the stories to showcase their abilities first as investigative reporters and second as multimodal storytellers. The stories made reference to the danger the reporters had put themselves in; the hidden nature of the information they had uncovered; and the effort they went to in analysing complex and hard-to-find data. The researcher also observed that the importance of the story topics was emphasised in order to justify the effort expended – what Dowling and Vogan (2015) pithily referred to as "Can we 'Snowfall' this?"

Reader Expectations

The researcher had preconceived ideas about the news organisations which influenced the way they read the stories. For example, the researcher assumed the BBC would make a point of showing neutrality while the *Washington Post* would favour US values at the liberal end of the political spectrum. Similarly, the researcher had expectations of professionalism from the *Guardian* and the BBC which influenced the level of trust placed in them. As a non-American, the researcher had no preconceptions about the *Tampa Bay Times*, and therefore did not approach the Scientology story through any

presumptions about the newspaper. Other readers might have different expectations and therefore react differently. What is described here is an awareness that most readers would have expectations of some kind based on their preconceptions about a news organisation.

Agenda in Evidence

Given that autoethnographic textual analysis involves a cognitive processing of the stories with critical thinking in mind, it is to be expected that the researcher observed agendas in each story. The question of 'why are they showing me this?' was applied to each mode. For example, the *New York Times* appeared to demonstrate anti-China undercurrents in its reporting on the New Silk Road. For instance it showed an image of a rusting tank with its barrel pointing towards China (implying threat) and made the point that the system used to control the railway was archaic (pre-1991). This sparked suspicion in the researcher's mind of a social, political, economic or cultural agenda that the situation represented in the story looked old fashioned and untrustworthy, rather than other frames such as it being a voyage of commerce and adventure. The researcher also noted a disparaging stance towards the Church of Scientology from the *Tampa Bay Times*, a political framing of the deepfake issue in the *Guardian*, and a lack of balance in the BBC's appraisal of China's Uighur camps.

This is not to say that the angles the stories took were unwelcome or had been misjudged by the newsroom. But a complex narrative that invites greater time, effort and commitment on the part of a reader may also receive more cognitive than heuristic processing and may attract the kind of critical thinking that literacy intends. For example, is the Church of Scientology really as untrustworthy as the story suggests and why is the *Tampa Bay Times* so harsh on it? Is the New Silk Road so lumbering and antique, and is there a political or economic benefit for *NYT* to frame China is such critical terms? Deepfakes are a typical example of the corruption of technology for nefarious purposes, so is this a case of the *Guardian* aiming to gain credibility as a serious news source by angling its story towards politics rather than entertainment?

Potter's second component is a personal locus which can override heuristic processing of news media and drive motivation to engage with the narrative and interpret the texts.

Interaction with Modes

The literacy skills of processing each individual mode scarcely need to be described. What is novel are the skills required to interact with the story. Scrolling brought rewards; the stories are created for this to happen. In the BBC story about Uighur camps in China, scrolling down an aerial view showed a building being constructed on empty land, giving the reader a sense of control. Interaction was not always successful, however. Clicking on images and anticipating a reward sometimes had no effect. The *Guardian's* deepfake story raised a flurry of boxes hiding an image; clicking removed them, only for them to pop back up: expectations of interactivity could be disappointed.

Reward and Pleasure

Further, the researcher observed that because they had clicked on an item, they felt more invested in the story. Linked to the apparent showcasing by the news outlet, an unexpected observation was that the researcher felt rewarded by the stories. There was a sense of pleasure at completion; in assembling the disparate parts into a whole; and at recognising tactics used to tell the story in a particular way. Pleasure was both an intended outcome for the reporters and sought out by the researcher through interaction with a story which is designed to be interacted with. Finally, there was pleasure in the movement of elements – a box inviting a click, a change of colour in a map, a counter on an audio clip showing seconds elapsed and remaining – which were more engaging than simple text and photography.

Shared reference points also elicited pleasure when they were identified, making the researcher feel a connection to the story. The New Silk Road through Kazakhstan may not be well known, but the cargo of Hewlett-Packard computers photographed is a reassuring reference point that helped the researcher connect with the story. The BBC report on Uighur re-education camps (new to the researcher) showed an aerial view of building on a football pitch (known to the researcher). The *NYT* reported Everest (which the researcher is unlikely to climb) in terms of the Taj Mahal and Burj Khalifa which are more easily accessible and understood reference points. Recognition of this tactic shifted engagement into a cognitive level; equally, such a tactic could be unconsciously appreciated by a reader, suggesting heuristic processing.

Potter's third component involves the competencies and skills demanded to process the media in its varied modes.

Connecting the Elements

Each shift between modes carries something from one to the next. At its simplest, the researcher carried a memory of a photograph when reading a caption that explained it; and then carried the insight from the caption back to the photograph to make meaning of it. The *Washington Post* story on climbing Everest was told through illustration interspersed with audio clips of people telling their experiences. The researcher saw these clips as a mechanism to introduce reality into a story otherwise told through illustration, making the final package more true-to-life.

Stories were set up to make connection easy: the researcher observed a flexible physical line linking each image of the New Silk Road to its place on a map; and in the story on climbing Everest, the three elements of a small schematic of the mountain on the left of the screen, the illustration of what can be seen at different elevations, and the text and audio describing it, were all connected by a line across the screen. This suggests that the reporters have given thought to how a reader would connect the elements. However, when two modes do not connect, it produces tension for a reader. The *NYT* story on the New Silk Road writes about heavy security for each shipment, but the photograph shows just one man in a hi-vis jacket and another on a bicycle – which scarcely registers as 'heavy' security – leading to a sense of manipulation. In the *Guardian* story on deepfakes, many of the images are non-political but the text refers primarily to the political threat, again leading to a sense of a mismatch

between the facts presented and the reporter's interpretation of their significance or the overall tenor of the story.

Potter's fourth component is how the information is processed in order to manufacture meaning or match it to existing opinions, worldviews or schema.

Interpretation and Making Meaning

The components of cognitive news literacy outlined by Potter do not extend to the interaction between modes, a cumulative *gestalt* effect of the way in which they work together. For example, the use of language indicated how the reporter hoped a reader would interpret the story. The *Washington Post* Mount Everest story refers to recent 'tragedies' in the standfirst to set the scene in terms of risk and danger; in the *Tampa Bay Times* the Church of Scientology is reported in terms such as 'secretive', 'control' and 'hostile' and suggestions of subterfuge without being clear what the subterfuge is; the BBC report on Uighur re-education camps used words such as 'watchtower' and 'razor wire' redolent of World War Two concentration camps, and ended with the words "history holds many troubling precedents about where such a project might end up." Given this, the researcher found themself interpreting non-text elements in these terms: as the russet of the illustration of the lower slopes of Everest gave way to cool blue, the researcher interpreted the colour as dangerous cold; the buildings highlighted on the map in the Scientology story took on a secretive tone; a photograph of a padlock on a traditional Chinese door was interpreted by the researcher as a metaphor for restriction and secrecy.

Conclusion

The starting premise for this study is that "each narrative medium requires a specific analytical approach to narrative structures and levels" (Onega and García Landa 2014, 2). This is particularly true for the varied possibilities in voice, point-of-view, media, modes and structure of multimodal longform. This article therefore attempts a specific analytical and narratological approach to the structure and modes of this form of digital journalism, in the form of autoethnographic textual analysis. The first two research questions have been answered; this section considers the larger question of the literacy that becomes visible when a reader 'joins the dots', as it is indicated by this new research method.

The method takes a cognitively aware, reflexive and narratological approach to the study of literacy in which narratology is "the organized and coherent analysis of narrative structures … showing up basic patterns and relations neglected in the more interpretive Anglo-American tradition" (Brooks 1984: xiii). Narratology is concerned with the way a narrative relates to the world of ideas that surrounds it (Darby 2001), and as such, the autoethnographic textual analysis set out here is a narratological exercise. Its advantage is that it allows for intense, focused and holistic study of navigation, interaction with individual modes in a multimodal story, combining these modes and interpreting their meaning, and critical thinking about agendas. This makes it apt to investigate literacy, which concerns the symbolic representation of culture,

ethics and knowledge, the skills to interpret media messages, and the institutional management of power which comes to those who can access and use knowledge (Livingstone 2004). Its contribution is to offer a method to join the dots allowing for consideration of the dots themselves, the way they join, and the picture of literacy they ultimately reveal.

The skills to navigate multimodal longform include such seemingly mundane tasks as scrolling and clicking, but also interaction with each mode, moving between them and carrying meaning across each move. Reading it requires the physical activity of moving the eye between modes and the hand in scrolling and clicking; the cognitive activity of understanding what is represented by each symbol – textual, graphic, video – at a simple level of literacy; and interpretive activity, exposing the new information to existing schema or 'worlds of ideas' (Darby, 2001). Digital multimodal longform is often designed for this to happen intuitively so that the story is easy to read. It is made easier by the linearity of scrolling but may be made more complex by peripheral media such as illustrations, interactive graphics and hyperlinks which offer distraction as much as enrichment. Debate continues over whether fragmented, hypertextual reading helps promote comprehension, critical literacy and choice (Burbules and Callister 2000; Reinking 1998) or lead to overload processing the information, confusion and frustration (Balcytiene 1999; Foltz 1996). The debate is spurious. If it is successful in the former, then strong literacy skills are in evidence; if it is not successful, then the literacy skills of either creator or consumer are in question.

Digital multimedia longform demands a reader to combine modes as it "converges the powerful elements of old media such as print and theater and honors the diverse forms of literacy that mark twenty-first-century audiences" (Dowling and Vogan 2015, 214). Coining the term "media grammars," Gumpert and Cathcart (1985) propose that there are important norms that are exclusive to each medium which frame people's media consciousness, and information processed through these norms shapes values and judgements. What makes multimodal longform distinct is a new grammar of *combination*. A narratological approach such as the one taken by this article would suggest that future literacy studies could concern themselves with what different modes do both singly and in conjunction: do they give flow to a story, continuity, exposition, information, context, explanation, illustration, and do they suggest manipulation, persuasion, emotion, influence?

A reader makes connections between disparate modes to construct a story and find meaning in it by interpreting the relationship between modes. The meaning of one may be altered when it is set alongside another, such as music or text influencing the meaning made in a photograph. "More than simply asking what modes or multimodal texts *are*, we need to be asking what multimodal texts *do* … to consider how multimodal texts work *intramodally* (how meaning is constructed *within* modes) as well as *intermodally* (how meaning is constructed *across* modes" (Serafini 2015, 413, his italics). The inter-connectivity and intra-connectivity of multimodal longform both merit study into how they produce meaning.

An overlooked aspect of literacy is the emotional impact of a text observed in this analysis, the sense of pleasure or disappointment it produces for the reader and how that impacts on its interpretation (see Wahl-Jorgensen 2013, 2020 for discussion of

emotionality in news). This study noted similarities between the product of multi-modal longform and the process of reading it. Two features of the product are pleasure when the reader agrees with the agenda, and disappointment when the reader does not. Similarly, the process of reading elicited pleasure when scrolling or clicking or combining elements delivered a reward, and disappointment when they do not. These emotional responses to both the process and the product affect engagement with the text and the literacy skills employed. This argues for a greater emphasis on emotionality as an aspect of journalism, as multimodal longform is often constructed to have a more powerful emotional impact than everyday news: "You have to make an effort to remember that the emotional engagement is achieved by the combination between a good story and the design of the experience that will be submitted to the audience" (Vázquez-Herrero, López-García, and Irigaray (2020), 36–37). Given the effort expended, the objective of multimodal longform is to engage the reader and make them more transported. Emotionality appears to be a key element of literacy as it impacts on critical thinking, and should be studied in future literacy studies (Orgeret 2020). Future studies might also interview the teams who create multimodal longform to identify what motivates them in their work and the importance they ascribe to factuality and emotionality, for example, or how they reconcile those two.

One limitation of this new method is the regular critique of any autoethnography – that of the limits of self-awareness. Wall (2008, 41) cites Clough (1998) that "no subject can be a fully self-identified, fully aware, or fully intentional author because unconscious desire makes fully intentional subjectivity impossible." Reflexivity is therefore a critical element of the approach. Yet, reflexivity as a critical constituent part of autoethnography also highlights the idiosyncrasy of any one individual researcher's reaction to a news item; other readers might react in a different way, based on their own life experiences, political stance, cultural background and so on. Yet even though their reactions vary, the process of reaction is consistent. The literacy revealed may be individual in its meaning, but consistent in its processes. The individualist nature of autoethnographic textual analysis make it suitable for initial forays into a subject such as literacy; greater complexity and nuance may be gained through focus-group discussions, involving the insights of everyday readers with those of more 'expert' scholarly researchers. Yet, autoethnographic textual analysis offers the benefit of a singularity of purpose and a level of cognitive processing and awareness which may be absent in focus groups, which argues for the two methods to be used in conjunction for future studies.

Finally, literacy also concerns production of media (Hobbs 2007). The skills required for journalists to produce multimodal narratives, where modes combine to have a particular effect or advance a particular agenda, find their mirror image in the effect on the reader. Media-literate readers should therefore be aware of the agendas of media producers. Identifying an agenda suggests successful critical thinking, which this study suggests merits further investigation as it relates to pleasure and disappointment. Media content is created to reward or please; sometimes it fails. The time-honoured uses and gratifications theory posits that it is consumed seeking reward (Katz, Blumler, and Gurevitch 1974); sometimes it is not successful, however. A consideration of a reader's expectations of pleasure and reward, and its counterweight of

disappointment, might therefore valuably contribute to theorising on literacy: what effect does emotionally pleasing sense of reward have on critical faculties, for instance;

Allied to this are expectations based on brand values, which also impact on critical thinking when engaging with any piece of multimodal longform or any media. A movie-goer likes the work of a certain star or director and anticipates the next film; a music fan looks forward to a new album by a singer; a reader returns to the same newspaper each day, all based on expectations of reward because of the values associated with the star, director, singer, or newspaper. Both brand values and the expectations of reward based on them form a vital part of literacy, drawing some of the dots which the reader anticipates will connect deliver a certain picture, and assemble them accordingly.

Disclosure Statement

No conflict of interest to declare.

Referen ces

Adams, T. E., C. Ellis, and S. Holman Jones. 2017 . "Autoethnography." In *The International Encyclopedia of Research Methods*. Wiley Online Library.

Ahva, L. 2013. "Public Journalism and Professional Reflexivity." *Journalism* 14 (6): 790–806.

Ashley, S., A. Maksl, and S. Craft. 2017. "News Media Literacy and Political Engagement: What's the Connection?" *Journal of Media Literacy Education* 9 (1): 79–98.

Balcytiene, A. 1999. "Exploring Individual Processes of Knowledge Construction with Hypertext." *Instructional Science* 27 (3-4): 303–328.

BBC. 2018. "China's Hidden Camps." Accessed May 15, 2021. https://www.bbc.co.uk/news/resources/idt-sh/China_hidden_camps.

Branch, J. 2012. "Snow Fall: The Avalanche at Tunnel Creek." Accessed May 21, 2021. https://www.nytimes.com/projects/2012/snow-fall/index.html#/?part=tunnel-creek.

Brooks, P. 1984. *Reading for the Plot: Design and Intervention in Narrative*. Cambridge, MA: Harvard University Press.

Burbules, N. C., and T. A. Callister. 2000. *Watch IT: The Risks and Promises of Information Technologies for Education*. Boulder, CO: Westview.

Cervetti, G., J. Damico, and P. Pearson. 2006. "Multiple Literacies, New Literacies, and Teacher Education." *Theory into Practice* 45 (4): 378–386.

Clough, P. T. 1998. *The End(s) of Ethnography*. New York: Peter Lang.

Coiro, J., and E. Dobler. 2007. "Exploring the Online Reading Comprehension Strategies Used by Sixth-Grade Skilled Readers to Search for and Locate Information on the Internet." *Reading Research Quarterly* 42 (2): 214–411.

Craft, S., S. Ashley, and A. Maksl. 2016. "Elements of News Literacy: A Focus Group Study of How Teenagers Define News and Why They Consume It." *Electronic News* 10 (3): 143–160.

Darby, D. 2001. "Form and Context: An Essay in the History of Narratology." *Poetics Today* 22 (4): 829–852.

Delamont, S. 2009. "The Only Honest Thing: Autoethnography, Reflexivity and Small Crises in Fieldwork." *Ethnography and Education* 4 (1): 51–63.

Dowling, D., and T. Vogan. 2015. "Can We 'Snowfall' This? Digital Longform and the Race for the Tablet Market." *Digital Journalism* 3 (2): 209–224.

Duncan, M. 2004. "Autoethnography: Critical Appreciation of an Emerging Art." *International Journal of Qualitative Methods* 3 (4): 28–39.

Dunham, R. S. 2020. *Multimedia Reporting: How Digital Tools Can Improve Journalism Storytelling*. Beijing: Springer.

Eason, D. 1984. "The New Journalism and the Image-World: Two Modes of Organizing Experience." *Critical Studies in Mass Communication* 1 (1): 51–65.

Eshet-Alkalai, Y. 2004. "Digital Literacy: A Conceptual Framework for Survival Skills in the Digital Era." *Journal of Educational Multimedia and Hypermedia* 13 (1): 93–106.

Ettema, J. 2005. "Crafting Cultural Resonance: Imaginative Power in Everyday Journalism." *Journalism* 6 (2): 131–152.

Foltz, P. W. 1996. "Comprehension, Coherence, and Strategies in Hypertext and Linear Text." In *Hypertext and Cognition*, edited by J.-F. Rouet, J. J. Levonen, A. Dillon, and R. J. Spiro, 109–136. Mahwah, NJ: Lawrence Erlbaum.

Giles, F., and G. Hitch. 2017. "Multimedia Features as 'Narra-Descriptive' Texts: Exploring the Relationship between Literary Journalism and Multimedia." *Literary Journalism Studies* 9 (2): 75–91.

Green, M. C., T. C. Brock, and G. F. Kaufman. 2004. "Understanding Media Enjoyment: The Role of Narrative Engagement into Narrative Worlds." *Communication Theory* 14 (4): 311–327.

Guardian. 2019. "The Rise of the Deep Fake and the Threat to Democracy." Accessed May 15, 2021. https://www.theguardian.com/technology/ng-interactive/2019/jun/22/the-rise-of-the-deepfake-and-the-threat-to-democracy.

Guillemette, L., and C. Lévesque. 2019. "Chapter 18: Narratology." In *An Introduction to Applied Semiotics*, edited by L. Hébert and J. Tabler. London: Routledge.

Gumpert, G., and R. Cathcart. 1985. "Media Grammars, Generations, and Media Gaps." *Critical Studies in Mass Communication* 2 (1): 23–35.

Hassan, S. 2020. "Digitality, Virtual Reality and the 'Empathy Machine'." *Digital Journalism* 8 (2): 195–212.

Haynes, K. 2012. "Reflexivity in Qualitative Research." In *Qualitative Organizational Research: Core Methods and Current Challenges*, edited by G. Symon and C. Cassell, 72–89. London: Sage.

Hiippala, T. 2017. "The Multimodality of Digital Longform Journalism." *Digital Journalism* 5 (4): 420–442.

Hobbs, R. 2007. "Approaches to Instruction and Teacher Education in Media Literacy." *Higher Education Research & Evaluation*, 58–64. https://unesdoc.unesco.org/ark:/48223/pf0000161133

Holman Jones, S. 2007. "Autoethnography." In *The Blackwell Encyclopedia of Sociology*. Wiley Online Library.

Horrigan, J. B. 2019. "Digital Readiness Gaps." Accessed June 15, 2021. https://www.pewresearch.org/internet/2016/09/20/digital-readiness-gaps/

Jacobson, S., J. Marino, and R. E. Gutsche. 2016. "The Digital Animation of Literary Journalism." *Journalism* 17 (4): 527–546.

Katz, E., J. G. Blumler, and M. Gurevitch. 1974. "Utilization of Mass Communication by the Individual." In *The Uses of Mass Communications: Current Perspectives on Gratifications Research*, edited by E. Katz, J. G. Blumler, and M. Gurevitch, 19–32. Beverly Hills, CA: Sage.

Knox, J. S. 2007. "Visual-Verbal Communication on Online Newspaper Home Pages." *Visual Communication* 6 (1): 19–53.

Kress, G. 2010. *Multimodality: A Social Semiotic Approach to Contemporary Communication*. New York, NY: Routledge.

Krippendorff, K. K. 2011. *Content Analysis: An Introduction to Its Methodology*. 2nd ed. Thousand Oaks, CA: Sage.

Livingstone, S. 2004. "Media Literacy and the Challenge of New Information and Communication Technologies." *The Communication Review* 7 (1): 3–14.

Maksl, A., S. Craft, S. Ashley, and D. Miller. 2017. "The Usefulness of a News Media Literacy Measure in Evaluating a News Literacy Curriculum." *Journalism & Mass Communication Educator* 72 (2): 228–241.

McAdams, M. 2015. "Multimedia Journalism." In *Ethics for Digital Journalists: Emerging Best Practices*, edited by L. Zion and D. Craig, 187–201. New York: Routledge.

Menke, M. 2019. "Multimedia Journalism." In *The International Encyclopedia of Journalism Studies*, edited by T. P. Vos and F. Hanusch, 1–9. London: John Wiley and Sons.

Mihailidis, P., and B. Thevenin. 2013. "Media Literacy as a Core Competency for Engaged Citizenship in Participatory Democracy." *American Behavioral Scientist* 57 (11): 1611–1622.

New London Group. 1996. "A Pedagogy of Multiliteracies: Designing Social Futures." *Harvard Educational Review* 66 (1): 60–92.

New York Times. 2013. "Riding the New Silk Road." Accessed May 15, 2021. https://www.nytimes.com/newsgraphics/2013/07/21/silk-road/index.html.

Onega, S., and J. A. García Landa. 2014. *Narratology: An Introduction*. Abingdon: Routledge.

Orgeret, K. S. 2020. "Discussing Emotions in Digital Journalism." *Digital Journalism* 8 (2): 292–297.

Palau-Sampio, D., and P. Sánchez-Garcia. 2020. "Digital Resources in the Current Journalistic Narrative: Uses and Limitations of Hypertext, Multimedia and Interactivity." *Communication & Society* 33 (2): 1–16.

Pilgrim, J, and C. Bledsoe. 2013. "The Application of Web 2.0 Tools in Literacy Education." In *Technological Tools for the Literacy Classroom*, edited by J. Whittingham, S. Huffman, W. Rickman, and C. Wiedmaier, 27–45. New York, NY: IGI Global.

Pilgrim, J., and E. E. Martinez. 2013. "Defining Literacy in the 21st Century: A Guide to Terminology and Skills." *Texas Journal of Literacy Education* 1 (1): 60–69.

Potter, W. J. 2004. *A Theory of Media Literacy: A Cognitive Approach*. Thousand Oaks, CA: Sage.

Reinking, D. 1998. "Introduction: Synthesizing Technological Transformations of Literacy in a Post-Typographic World." In *Handbook of Literacy and Technology: Transformation in a Post-Typographic World*, edited by D. Reinking, M. C. McKenna, L. D. Labbo, and R. D. Kieffer, xi–xxx. Mahwah, NJ: Lawrence Erlbaum.

Sánchez Laws, A. L. 2020. "Can Immersive Journalism Enhance Empathy?" *Digital Journalism* 8 (2): 213–228.

Serafini, F. 2014. *Reading the Visual: An Introduction to Teaching Multimodal Literacy*. New York, NY: Teachers College Press.

Serafini, F. 2015. "Multimodal Literacy: From Theories to Practice." *Language Arts* 92 (6): 412–423.

Siegle, R. 1986. *The Politics of Reflexivity: Narrative and the Constitutive Poetics of Culture*. Baltimore, MD: Johns Hopkins University Press.

Silverblatt, A. A. Smith, D. Miller, J. Smith, and N. Brown. 2014. *Media Literacy: Keys to Interpreting Media Messages*. Santa Barbara, CA: Praeger.

SmartOcto. 2020. "5 Great Examples of Innovative Longform Journalism Online." Accessed May 15, 2021. https://smartocto.com/blog/5-great-examples-innovative-longform-journalism-online/.

Sparkes, A. C. 2002. "Autoethnography: Self-Indulgence or Something More?" In *Ethnographically Speaking*, edited by A. P. Bochner and C. Ellis, 209–232. Walnut Creek, CA: AltaMira.

Tampa Bay Times. 2019. "Clear Takeover: How Scientology Doubled Its Downtown Footprint in 3 Years." Accessed May 15, 2021. https://projects.tampabay.com/projects/2019/investigations/scientology-clearwater-real-estate/

Thoman, E., and T. Jolls. 2005. "Literacy for the 21st Century: An Overview and Orientation Guide to Media Literacy Education." CML Media Lit Kit. Accessed June 15, 2021. athttp://www.medialit.org/sites/default/files/01_MLKorientation.pdf

Thomas, J. 1993. *Doing Critical Ethnography*. Newbury Park: Sage.

Tversky, A., and D. Kahneman. 1974. "Judgment under Uncertainty: Heuristics and Biases." *Science* 185 (4157): 1124–1131.

Vázquez-Herrero, J. X. López-García, and F. Irigaray. 2020. "The Technology-Led Narrative Turn." In *Journalistic Metamorphosis*, edited by J. Vázquez-Herrero, S. Direito-Rebollal, A. Silva-Rodríguez and X. López-García, 29–40. Switzerland: Springer.

Vraga, E. K., M. Tully, A. Maksl, S. Craft, and S. Ashley. 2021. "Theorizing News Literacy Behaviors." *Communication Theory* 31 (1): 1–21.

Wahl-Jorgensen, K. 2013. "Subjectivity and Story-Telling in Journalism: Examining Expressions of Affect, Judgment and Appreciation in Pulitzer Prize-Winning Stories." *Journalism Studies* 14 (3): 305–320.

Wahl-Jorgensen, K. 2020. "An Emotional Turn in Journalism Studies?" *Digital Journalism* 8 (2): 175–194.

Wall, S. 2008. "Easier Said than Done: Writing an Autoethnography." *International Journal of Qualitative Methods* 7 (1): 38–53.

Wallis, R., and D. Buckingham. 2019. "Media Literacy: The UK's Undead Cultural Policy." *International Journal of Cultural Policy* 25 (2): 188–203.

Ward, S. J. A. 2004. *The Invention of Journalism Ethics. The Path to Objectivity and Beyond*. Montreal: McGill-Queen's University Press.

Washington Post. 2016. "Scaling Everest." Accessed May 15, 2021. https://www.washingtonpost.com/graphics/world/scaling-everest/.

Weir, D., and D. Clarke. 2018. "What Makes Autoethnographic Analysis Authentic?" In *Ethnographic Research and Analysis: Anxiety, Identity and Self*, edited by T. Vine, J. Clark, S. Richards and D. Weir, 127–154. London: Palgrave Macmillan.

Wu, J., and Y. M. Wang. 2011. "Unpacking New Media Literacy." *Journal of Systemics, Cybernetics and Informatics* 9 (2): 84–88.