
Book Section:
Parry, K orcid.org/0000-0003-3654-6489 (2020) Quantitative Content Analysis of the Visual. In: Pauwels, L and Mannay, D, (eds.) The SAGE Handbook of Visual Research Methods. SAGE , pp. 352-266. ISBN 9781473978003

https://doi.org/10.4135/9781526417015

26

Quantitative Content Analysis of the Visual

Katy Parry

INTRODUCTION

Twenty years have now passed since Hansen et al. (1998: 189) described visual analysis as the 'poor relation' in mass communication research. Within the traditions of social science research, the 'problematic nature of visual representation' (190) is cited as a possible reason for this long-standing neglect; with the rational word privileged over the emotive image. In other words, the particular meaning-making potentials of visual images are not considered to be captured effectively by tools of analysis that treat them primarily as objective or transparent 'windows' into reality. If we have learned anything from semiotics, cultural studies, feminist theory and the 'affective turn', surely it is that images are polysemic, symbolic, performative and emotive. Why, then, would we want to embrace the positivist tradition of quantitative content analysis when examining visual materials?

This chapter presents an overview of quantitative visual content analysis as a method, how it can support scholars in answering certain research problems or questions and how to circumvent its shortcomings through robust design and the use of complementary combined methods. I hope to show that content analysis helps us to examine important questions of ideological influence and the ways in which particular versions of reality are constructed and fashioned over time and across a variety of media. I explore the appropriate applications for

such a systematic approach to a diverse range of visual media content, illustrated with examples from the methods literature and archetypal studies.

To demonstrate how to design and apply this method, I present a codebook from my own study on the United Kingdom (UK) press photographic depiction of the Iraq War in 2003. In addition to showing how the coding system works with categories adapted from past studies, the aim here is to illustrate how fragmentary bits of information can be cross-tabulated in order to interrogate the portrayal of people and events: in this case, the depiction of civilian casualties during a time of war. Finally I turn to how the digital age has reconfigured both what we mean by 'content' and the kinds of recordable features and measures that intrigue researchers of the visual. In the age of 'big data' what might count as meaningful communicative content? This chapter is not intended as a guide to complex statistical analysis of aggregate data, but rather to show how those who take visual images seriously can benefit from integrating quantitative content analysis into their repertoire of methods.

WHAT IS QUANTITATIVE CONTENT ANALYSIS?

Before turning to visuals in more detail, it is worth setting out what is meant by quantitative content analysis. Perhaps one of the most oft-cited definitions is Bernard Berelson's: 'Content analysis is a research technique for the objective, systematic, and quantitative description of the manifest content of communication' (1952: 18). Whilst some might take issue with whether any technique is entirely 'objective', the important take-away words here are 'systematic' and 'manifest content'. Content analysis has to follow a series of steps in its design to ensure it is measuring relevant content in a consistent manner, so a systematic approach is essential. Berelson's 'quantitative description' means that the content is summarized numerically (usually presented in tables of results): it is reported in aggregated

form, rather than as detailed discursive description. Such a method therefore allows for the 'mapping [of] key characteristics of large bodies of text' and 'follows clearly articulated rules and procedures' (Hansen et al., 1998: 123-124). And where research of the *visual* is central to the design, 'text' expands to comprise visual content rather than the typical linguistic elements.

For Asa A. Berger, content analysis 'is a research technique that is based on measuring the amount of something (violence, negative portrayals of women, or whatever) in a representative sampling of some mass-mediated popular art form' (Berger, 1998: 23). Again this is about counting or measuring something manifest, but Berger points to just how varied this content might be, and hints at another important stage in the process – sampling of *representative* materials. Even with a method suited to large amounts of media materials, a careful process of selection is often required: enough to defend its representativeness whilst ensuring feasibility.

In one of the major content analysis guidebooks, Neuendorf (2002: 4) readily admits confusion within the academic literature as what counts as content analysis – its variety of applications from across disciplinary boundaries accounts for some lack of clarity here. A particular murky area when it comes to visual content is the distinction between quantitative and qualitative content analysis, and Berger's 'negative portrayals of women' above would be a case in point. At what point are we in fact using qualitative judgements in designing and conducting visual content analysis? I will return to this distinction below.

A more recent and comprehensive definition from Riffe et al.'s *Analyzing Media Messages* (2014) moves beyond the scientific processes of the method to take account of why we might want to conduct a content analysis:

> Quantitative content analysis is the systematic and replicable examination of symbols of communication, which have been assigned numeric values according to valid measurement rules, and the analysis of relationships involving those values using statistical methods, to describe the communication, draw inferences about its meaning, or infer from the communication to its context, both of production and consumption. (Riffe et al., 2014: 19)

It is not simply about being able to describe the manifest content in summarized form based on a rigorous scientific process, but crucially it is about making inferences about communicative significance and interconnections with other sites of meaning (context, production and consumption). 'Replicability' is another key word here: the idea being that if another researcher conducted a study using the same codebook on identical material, they would produce the same results. Importantly this also means that the researcher can conduct comparative or longitudinal studies using an equivalent codebook on different cases with an assured sense of valid comparison.

There are a number of key methods chapters for guidance when it comes to visual materials (Bell, 2001; Rose, 2012; Bock et al., 2011) but despite the presence of visual content in many forms of media, images remain on the periphery when it comes to quantitative methods handbooks. For example, in the third edition of Riffe et al.'s *Analyzing media messages: using quantitative content analysis in research* (2014), the index refers to only three pages for 'visual communication'. But a closer read shows that the third edition has actually been revised to further integrate visual content and in fact this volume offers a useful introduction to the method and especially for advertising and ethnicity.

In sum, content analysis is an excellent method for telling us 'what is there' in a large corpus of media content. But in order to both devise meaningful categories in the research design and to enable interpretation of why such resulting frequencies or omissions are

culturally or politically significant, the researcher needs to draw upon in-depth contextual understanding and theoretical foundations.

QUANTITATIVE OR QUALITATIVE? THE BENEITS OF A 'LONG SOAK' IN YOUR DATA

In the introduction to *Paper Voices* (1975), Stuart Hall describes their project's methods for analysing the press in its role as a 'social educator', mixing content analysis with literary-critical, linguistic and stylistic methods:

> Both methods are based on a long preliminary soak, a submission by the analyst to the mass of his [or her] material: where they differ is that content analysis uses this process of soaking oneself to define the categories and build a code (based on an intuitive sense of where the main clusters occur), whereas literary, stylistic and linguistic analysis uses the preliminary reading to select representative examples which can be more intensively analysed. (Hall, 1975: 15)

Hall actually stresses some similarities in what are often posited as contradictory methods; both benefit from a 'long preliminary soak' by the researcher, and they both 'employ evidence' (ibid). Hall's call to the researcher to 'submit' holistically to the material is perhaps influenced by Siegfried Kracauer's attack on the fragmentary character of quantitative content analysis, and his belief that the text is approached as a 'meaningful whole'.

> Most communications are not so much fixed entities as ambivalent challenges. They challenge the reader or the analyst to absorb them and react to them. Only in approaching these wholes with his whole being will the analyst be able both to discover and determine their meaning – or one of their meanings – and thus help them to fulfil themselves. (Kracauer [1953] cited in Larsen, 1991: 122)

Kracauer's *qualitative content analysis* allows for the recognition of various meanings in one text, of various discourses which may be both surface (manifest) and hidden (latent). The notion of 'communications' as 'ambivalent challenges' acknowledges the unfixed nature of meaning in various types of text and again raises the issue of what the analyst can discover through quantitative content analysis alone.

ADDRESSING THE LIMITATIONS OF CONTENT ANALYSIS FOR VISUAL MATERIALS

When researching images it sometimes feels like researchers of the visual start from a position of deficiency or defensiveness, accepting that images are 'problematic', that they lack a recognized syntax, that they are interpreted differently by different people (polysemic), that for moving images in particular there is no simple way to decide how to splice each frame or scene in terms of its meaningful content. In the case of visual content analysis this can be coupled with a recognition of the method's limitations as a rather 'blunt instrument' (Hansen et al., 1998: 190). Content analysis can be limited in nature, due to the reduction of complex information into numerical properties, which, on its own, can tell the analyst very little about meaning or intentions of the content. Content analysis can reveal the frequency and prominence of certain content of interest, but not 'why-the-content-is-like-that' (Berelson, 1952: 16), or the socio-cultural significance indicated by the aggregated results.

The positivist slant of content analysis and its emphasis on objective 'manifest' content raises concerns when photographic images are the primary focus of the study. Berelson writes that 'content analysis must deal with relatively denotative communication materials and not with relatively connotative materials' (1952: 19), echoing Kracauer's denotative and connotative distinctions. Yet both words and photographs arguably contain both levels of communication. To switch the emphasis, I would argue that it is not the

limitations of the materials, in terms of clarity of meaning, which is the primary concern here, but the limitations of the method that needs to be addressed, and so enriched with another technique that is apt to deal with meanings.

However, one purported limitation of content analysis is perhaps warranted: 'It is very hard to evoke the mood or the affect of an image through codes' (Rose, 2012: 102). Social scientific methods are not well suited to capture the expressive or emotive qualities of visual images. Likewise other developments in visual communication analysis have tended to over-emphasize the syntax or grammar of images, attempting to understand them through fixed rules rather than an emotional connection. Media scholars interested in visual representation are likely to focus on issues of class, gender or ethnicity through stereotypical depictions rather than their emotive qualities. Supplemented with a rhetorical, discourse or framing analysis that aims to explore the thematic, strategic and persuasive elements of a visual image, content analysis provides the building blocks for a more sophisticated approach.

WHAT TYPES OF VISUAL MEDIA CONTENT HAVE BEEN EXAMINED USING QUANTITIATIVE CONTENT ANALYSIS

Despite some of the concerns raised above, content analysis is a widely used method across fields of media research, with visuals considered alongside textual and narrative aspects, if not taken as the primary analytical focus. Table 1 summarizes some of the classic or recent studies of visual content analysis in order to highlight the variety of applications and the kinds of concepts or theories being tested.

Table 1: Studies conducting some form of quantitative content analysis of visual materials

Media form	Example study subject and citation	Theory/concept
Advertising	*Gender Advertisements*, (Goffman, 1979); see also (Bell and Milic, 2002)	Gender relations and stereotyping, rituals of subordination, gestures
Magazines	*Reading National Geographic*, (Lutz and Collins, 1993) Visually framing the Iraq war in *TIME*, *Newsweek*, and *U.S. News & World Report* (Schwalbe, 2013)	Orientalism, exoticism, power, race, colonial history War photography, graphic imagery, gender, visual framing
Film	*The White Savior Film*, (Hughey, 2014)	'savior' trope and whiteness, masculinity
Television news	*The Uncensored War: The Media and Vietnam*, (Hallin, 1986)	oppositional media, spheres of consensus, deviance and legitimacy
Tourism publicity materials	Feminity in Macau's tourism adverts (Sun, 2017)	Femininity, post-colonialism, labour roles, consumerism
Multimedia packages and online news galleries	Multimedia packages on New York Times website, (Jacobson, 2012) Photos of the Day news galleries, (Midberry, 2017)	'transcoding' of journalism in digital environment Newsworthiness, 'shared commonality'
Social media	YouTube videos (van Zoonen et al., 2010) Profile pictures on Facebook (Emmons and Mocarski, 2014) 'Selfies' on Instagram (Veum and Undrum, 2017)	Citizenship, video activism Branding, gender, athletes Self-representation, gender, global genres
Photo libraries/ archives	Afghan women in AP (Associated Press) photographs, (Fahmy, 2004)	visual subordination and stereotypes, Orientalism
Newspaper photography	Civilian casualty photographs during the Iraq war (Parry, 2012)	Bearing witness, human cost of war, graphic imagery
Political cartoons	Political cartoons before and after Pearl Harbor in Chicago newspapers, (Vultee, 2007)	Visual rhetoric, foreign policy, patriotism, 'rally effect', cartoon themes
Political campaign material	Online political posters on Facebook, (Lee and Campbell, 2016)	Strategic roles and themes for political persuasion and mobilization

Table 1 presents a number of studies, and the sheer range quickly becomes apparent. The more recent studies demonstrate the continuing utility of content analysis as a method, although some studies do not necessarily promote themselves in this way. Perhaps two of the

most influential studies are Erving Goffman's *Gender Advertisements* (1979) and Catherine A. Lutz and Jane L. Collins' *Reading National Geographic* (1993). Whilst Goffman's characteristics of gender stereotyping, such as 'the feminine touch', have remained prominent in later studies on advertising, he has come under criticism for the rather opaque selection process of the chosen adverts (see Bell and Milic, 2002).

But Lutz and Collins (1993) present a relatively clear rationale, selection process and coding strategy for their study of *National Geographic* magazine, covering from 1950 to 1986. With an interest in how 'cultural difference' between Westerners and non-Westerners was constructed in photographs from around the world, the content analysis revealed which regions of the world were favoured in reporting and how this shifted over time. Using variables such as 'world location', 'camera gaze of person photographed', 'urban versus rural setting', 'male/female nudity' and 'vantage' (285), Lutz and Collins could start to build a picture of the patterns of representations, and especially how troubling concepts of race and gender were perpetuated. The authors' own analysis is also enriched by the inclusion of institutional and audience perspectives gained through interviews with magazine editors and readers.

A number of the later studies in Table 1 reflect the development of internet technologies, integrating extra-textual features such as hypertext links and audience feedback mechanisms in additional to characteristics of the content itself (narrative, expressions). For example, Susan Jacobson's (2012) study of nytimes.com charts how the 'non-journalistic voice' and 'first-person experience' have become more prominent through the increasing use of photo slideshows and videos over an eight year period. I will return to studies concerned with visual content in the digital environment later. Now that we have a sense of the various possible uses for content analysis, the next section turns to practical guidelines on how to apply this method.

HOW TO GO ABOUT DESIGNING A CONTENT ANALYSIS OF VISUALS

In keeping with its scientific traditions there are a standard set of stages required in devising a content analysis. The next section introduces each of the stages: theoretical framework and concept definitions; devising categories to measure meaningful dimensions; sampling of media materials; codebook construction; coding and reporting of results.

Theoretical framework and concept definitions

Whether primarily interested in visual content or multiple modes of communication, it is important to have a good sense of supporting theories or key concepts. For example, in a study on post-colonial Macau's tourism publications, Zhen Sun (2017: 2625) examines 'to what extent patriarchal ideology, stereotypes, and the male view have been ingrained in tourism practices and advertising' and so situates her research within theories of how gender is constructed in advertising in general terms and, more specifically, in relation to the 'forces of globalism and neoliberalism underlying the quick expansion of Macau's casino industry' (ibid).

Sun therefore builds her theoretical framework and defines concepts in relation to literature about gender stereotypes in advertising, patriarchal divisions of labour in tourism, post-feminism, neoliberalism and post-colonialism. This then feeds into the construction of her hypotheses and the codebook that is able to capture the kinds of depicted roles for men and women (massage therapist, chef, or shopper) and the relationships between them (co-workers, romantic couple). Sun finds that women are still often depicted as models, in 'decorative' mode, but it is also interesting to see how the casinos hoped to appeal to female friends as well as heterosexual couples, depicting a luxurious and safe space for shopping

together. In this way the author notes a 'feminization of gambling' in the patterns of gender representation, as advertisers 'place women in nearly all kinds of consumption spaces' (2641).

As a content analyst, once you define the core concepts, you need to ensure that the material selected and the characteristics you are quantifying will enable you to measure such concepts. In other words, do you have the right empirical indicators for the concepts you are hoping to test? There is a danger in devising a content analysis that the researcher wants to count everything. This can be time-consuming and unnecessary. Ensure that you are capturing relevant and meaningful characteristics or variables by connecting your key concepts or theories to your categories for analysis.

Categories, measures, variables, or values?

Different uses of these terms can be quite confusing, especially where variables and values are sometimes used inconsistently. For our purposes *variables* refers to the dimensions or features you want to record for each image/ item; whilst *values* are the various options you choose from within the variable category. Stepping back slightly, the first step is actually to decide what counts as your primary unit of analysis. It could be a single photograph, but it could be an advert which contains various visual and textual elements. For each unit or item, the list of variables represents the aspects of communicative information you hope to capture about its form and content, guided by your theoretical understanding. The values are the list of possibilities for each variable and will need to be decided upon before coding takes place.

As Bell writes, the values should be '*mutually exclusive and exhaustive'* (2001: 16) so that the coder can select one without too much ambiguity (and this is why there is usually an 'other' option in the value list for unexpected content). The values are often numbered so that coding decisions can be made quickly, and entered into a suitable spreadsheet or database (Excel, SPSS, Access). So how extensive are the 'best' lists of values? More detail is not

necessarily a good idea. To ensure variables and values are analytically meaningful, content analysts often draw upon past studies in addition to their own theorized concepts.

Sampling

Content analysis is useful for analysing a large amount of material, but a process of selection and sampling is likely. First, there will be a certain time period covered – this can be relatively clear-cut due to the nature of your study – an election campaign period, or the run of a TV series. But even here, a rationale is required. For example, modern wars and military interventions do not often have a clear beginning or end, and the time-period of the selected media material will affect the findings. Do you include the critical debate phase in the run-up to the war, or the difficult period of post-conflict transition following military operations?

Second, there is the sampling of chosen media materials: do you include television, print or online media? Which social media platform is most relevant to your study? Which newspapers best represent the spread of political perspectives and public debate? Third, there is the selection process *within* the corpus of chosen media materials: front pages or the whole magazine? Articles which mention a keyword within the headline, or first paragraph? Do you analyze every single episode broadcast? YouTube videos tagged with the keyword *and* viewed a certain number of times? Image-tweets accompanied by certain hashtags? Even where visual content alone is analyzed, linguistic content is often used in the selection process (tags, hashtags, keywords), meaning that the image-word relationship is difficult to avoid entirely.

There are a number of sampling strategies designed to select a portion of the material whilst ensuring it remains representative and relevant to your research. This might be arranging your data by date and then selecting every *n*th image (for example, every tenth image in a corpus of 1000 images would give you a sample of 100 images to analyze). Rather

than this systematic approach, random sampling is where you assign a number to each image and then randomly select the identifying number – software programmes such as Excel can do this for you. For more details on sampling methods see chapter six in Krippendorff (2004).

Codebook construction

As discussed above, deciding on your sample, unit of analysis, variables and values provides the foundations for operationalizing your content analysis study. Once you have a working idea of the sampled corpus and components you intend to record, it is essential to conduct a pilot study to test your coding categories. Testing your approach in this way reveals where omissions, ambiguities or overlaps occur. Do not forget that once coding starts 'for real', the coding strategy needs to be applied consistently, because you cannot make significant adjustments to the codebook once underway. Given the particular qualities of visual materials, it might be impossible to eliminate all ambiguities and this is where detailed instructions to coders and 'coding conventions' are useful – even if there is in reality only a single coder. The aim here is to create a set of instructions and precedents so that another 'imaginary' coder would be capable of following the coding strategy, in addition to building notes for observations which will aid later analysis.

Coding and reporting of results

As already mentioned, the coding stage often utilizes a spreadsheet or database package which facilitates coding and the later statistical analysis. With the codebook template set up and tested, the coding can begin. Along with representativeness and valid measurement rules, content analysis also needs to demonstrate its reliability. Put simply, can the values be assigned in a consistent manner? If there is more than one coder, each should be able to make the same judgement in coding decisions, possibly following some training to ensure the variables and values are fully understood: this is known as 'inter-coder reliability'. This

should be tested throughout the coding period; for example, all coders coding a set of the same items at the beginning, middle and end of the research task. If there is a single coder, their consistency also requires checking, coding a sample of the same items at different intervals to ensure constancy over time: this is known as 'intra-coder' reliability (for further details, see Bell, 2001: 21-4).

Aggregated results are generally presented in tables with both numbers and percentages given where appropriate. But be careful of assuming what the results reveal: careful explanation and interpretation of what the numbers *mean* and how they relate to the theoretical concepts underpinning the research requires expertise and contextual knowledge. As suggested earlier, supplementing statistical data with more detailed readings of key examples helps to make a convincing case when discussing the cultural significance of the visual materials.

CASE STUDY: PHOTOJOURNALISTIC DEPICTIONS OF THE 2003 IRAQ INVASION IN THE UK

In order to illustrate the considerations involved in codebook design and coding strategies, this section summarizes the variables used in the quantitative part of my study on the UK press photographic coverage of the 2003 US-led Iraq invasion. The war was hugely controversial and intensely covered in the news media. The study was designed to examine how certain realities and narratives were visually constituted in the photographic coverage of the war, in order to evaluate the degree to which the Coalition's justifications for the war were supported or countered. For instance, the rhetoric of 'humanitarian warfare' (of bringing liberation, democratization and human rights to the Iraqi people) was widely advanced by the US-led Coalition and its supporters as a legitimating reason for war. How might the images

of Iraqi civilians, receiving aid, greeting soldiers, or as hostile to the invasion, work to support or undermine this rationale?

The study includes seven national newspapers in the UK (*Sun*, *Mirror*, *Mail*, *Independent*, *Guardian*, *Times* and *Telegraph* – inclusive of their Sunday equivalents). The newspaper sample encompasses the diverse nature of the national press in the UK, comprising a mix of quality (broadsheet), mid-market and tabloid titles, with a range of positions of editorial support and opposition towards the war. The primary unit of analysis is the photograph, with the caption and headline also included where relevant. I take account of *all* Iraq-related photographs in the surveyed sample, providing 4,389 images during the 33-day period that reflect the 'major combat operations' period of the war (17 March–18 April 2003). On average, 19 photographs *per newspaper* appeared daily across the Iraq news pages. The quantitative analysis was designed to integrate semiotic features and is supported with a visual framing analysis, but for the purposes of this chapter I focus on the coding strategy for manifest visual elements related to both format and composition (for example, photo size on page; camera angle and distance; number/age/gender of subjects) and the surrounding text (headline/caption).

Coding Strategy for Iraq Invasion Study: Variables employed in MS Access Database

Below is an annotated version of the codebook, showing how the content analysis integrated measures from existing visual studies to support its design.

1. **Newspaper** (such as *Guardian*, *Times*)
2. **Date**
3. **Page**
4. **Size** Values: Small, Medium, Large, Extra Large

Additional Notes on Size: The size coding depends on the proportion of the page: over half a page is extra large; between a quarter and half page would be 'large'; less than a quarter but

more than an eighth is ‘medium’; and below that ‘small’. This is coded by sight, otherwise it would involve measuring and comparisons (which is considered too time consuming). The size of the photograph is understood as an indication of the news story’s salience or importance.

5. **Colour?:** (colour vs. black and white) See Lutz and Collins (1993: 93-4) on possible associations with the use of colour).

6. **Top of page:** (does any part of the image appear on the top half of the page?) Using both this control, and through coding the order of photographs on the same page, a certain hierarchy of types of image can be ascertained.

7. **Relevant Headline?** (free text box)

8. **Caption?** (free text box)

9. **Location:** (such as Iraq, Kuwait, chosen from a list of values) This is based on the amount of detail given in the photo, caption or headline – code as ‘unspecified’ if not apparent – do not spend time checking other sources or text to confirm location.

10. **Graphic Nature?**: In most cases the ‘graphic nature’ measure will have a ‘not applicable’ (‘n/a’) value coding. However, where casualties are depicted, this variable provides further information as to the visual impact of the photograph.

Not applicable
injury not obvious
face shown injured
face hidden injured
death see face no blood
death see face with blood
death, face obscured
death disembodied
metaphorical
mourning
coffins

Additional Notes on Graphic Nature: Shahira Fahmy adopts Potter and Smith’s ‘distance’ and ‘physical alteration’ determinants for identifying graphic imagery:

> A close-up image of a violent act is more graphic than a long shot of that same act […]. Second is the degree of physical alteration to the victim. For example, in portraying a body, in a nongraphic image the victim may lie with his or her eyes shut; in a graphic image, the victim may be lying in a pool of blood. (Fahmy, 2005: 148)

I add that further factors are crucial in determining the 'graphicness' of press photos; namely, the degree of framing or cropping, so that we might only see body parts rather than the whole body (I have labelled this 'disembodied' in the value list), and whether or not we see the face of the victim (see Hallin's detailed appendix for content analysis variables applied to coverage of the Vietnam war (1986: 260)). The printed *size* of the image and use of *colour* also has an impact on 'graphicness' (see measures 4. and 5.).

The next part of the coding sheet concentrates on **compositional features**.

11. **No. of subjects** (actors) (values count up to five – after that code 'many')
12. **Gender** (NA/male/female/mixed)
13. **Age** (NA/ child/ adult/ mixed)
14. **Distance?** Distance for central focus of the image?

Additional Notes on Distance: Here I use the typology developed by Kress and van Leeuwen (1996).

> Even though distance is, strictly speaking, a continuum, […] the close shot (or 'close-up') shows head and shoulders of the subject, and the very close shot ('extreme close-up', 'big close-up') anything less than that. The medium close shot cuts off the subject approximately at the waist, the medium shot approximately at the knees. The medium long shot shows the full figure. In the long shot the human figure occupies about half the height of the frame, and the

very long shot is anything ‘wider’ than that. (Kress and Van Leeuwen, 1996: 130)

Kress and van Leeuwen relate these distinctions to social and personal relations, noting how ‘patterns of distance can be conventional in visual genres’ (132).

15. Eye contact? **Direction of gaze** (only code for single dominant subject or central focus).

Direction of gaze
Not applicable
Indirect: left
Indirect: right
Direct: at reader
Indirect: other people responsive
Indirect: other people not responsive
Indirect: at object
Indirect: oneself
Indirect: downwards
Indirect: middle distance
Indirect: three-quarter gaze
Hidden by mask or glasses
Not discernable

16. **Angle?** ‘Not Applicable’; ‘High direct’; ‘Medium direct’; ‘Low direct’; ‘High oblique’; ‘Medium oblique’; ‘Low oblique’. (Kress and Van Leeuwen, 1996: 146-154; Hansen et al., 1998: 207-208)

Additional Notes on Gaze: Visual analysts write that certain relationships can be constructed through direct eye contact between the photographic subject and the photographer/ camera/ implied viewer. A subject’s direct gaze can be an invitation to make an emotional contact or alternatively it can be confrontational or accusatory (Lutz and Collins, 1993: 176, 197). Along with the distance and angle, the direction of the gaze can demand something of the viewer, creating ‘a visual form of direct address’ demanding ‘that the viewer enter into some kind of imaginary relation with him or her’ (Kress and Van Leeuwen, 1996: 122). Indirect

gazes can signify other associations: 'For politicians the three-quarter gaze is more common: a gaze that soars rather than confronts, suggesting instead of the relation to the viewer, to the present, the more ennobling abstract relation to the future' (Sontag, 1979: 38).

17. **Photo Subject**:

Main Heading > Sub Category

*Code for **more than one subject** if necessary*

Example values *(full list not included)*:

#	Main Category	Sub Category
001	Political People	Tony Blair
090	Military People	Coalition military: Tommy Franks
140	Civilians	Iraqi civilians
163	Other People	Aid/NGO/ Red Cross worker
210	Military Action	Armoured vehicle/tank
230	Humanitarian	Aid delivery
283	Destruction	Post-missile rubble

Notes on Photo Subject: For the purposes of this study, the codebook enabled the recording of multiple 'subjects' for each image, arranged under main category headings, so that the later analysis could be conducted either at the level of general category heading or the more specific sub-category. The 'subject' variable can also be cross-tabulated with other variables or measures to isolate photographs employing certain compositional features. For example, one might want to study how the presence of soldiers and weaponry is shown to directly impact on Iraqi civilians. How is the devastation/progression of the military campaign greeted by Iraqi people? The dynamic nature of the 'subject' classification allows the coder to merge these main subject categories to isolate photographs which, for example, depict both soldiers and Iraqis to examine how these encounters are depicted.

Similar to Lutz and Collins' (1993) study, the idea here is not only to present tables of aggregated numerical data, but to cross-tabulate the various aspects to demonstrate the

patterns of visualization, and also to return to the individual photographs for further analysis in line with the theoretical underpinnings driving the research.

Example from findings: Civilian casualties

To illustrate just one of the ways the content analysis was used to explore the favoured depiction of the war across the range of newspapers, I present the results pertaining to 'civilian casualties' (see Parry (2012) for the full discussion of these findings).

On average, across all newspaper titles, just seven per cent of the newspapers' Iraq-related photographs dealt with the issue of non-combatant casualties, including non-Iraqi and journalist casualties (*n*=309). So straight away, in terms of how the war's destructive consequences for local civilians were represented, we can see that it is a relatively small amount of visual coverage. However, within this coverage, it is also possible to observe that children featured highly in this aspect of coverage: 23.6% of photographs in this category pictured children *without* adults (*n*=73), and a further 21.0% depicted children with adults (*n*=65) – altogether, almost half the images in this category featured children (44.6%; *n*=138). This would suggest that where civilian casualty imagery appeared, there was a substantial focus on the suffering of the most vulnerable and innocent, with the potential to rupture the narrative of a moral purpose.

There are also substantial differences *between* the newspapers. This is most starkly observed in the two tabloid titles. The anti-war tabloid, the *Mirror*, prints the greatest proportion of casualty-related images (12.1 per cent, *n*=92) while the *Sun* has the least (2.6 per cent, *n*=15). Whilst the *Mirror* in particular refused to shy away from directly representing of the human cost of the conflict, the *Sun* newspaper is also exceptional in its near disappearance of injury and death.

Limiting the findings to photographic depictions of *Iraqi* civilian injury, grief or death, (i.e. excluding journalists and non-Iraqi civilians), Table 2 details the number of images with graphic content in each newspaper, drawing on Daniel Hallin's distinctions in his Vietnam study (1986: 260).

Graphic Nature/ Newspaper (n)	**Mirror**	**Indepdt**	**Guardn**	**Mail**	**Times**	**Telegph**	**Sun**
face shown injured	37	17	11	17	13	11	1
face hidden injured	1						
mourning	8	5	8	3	2	2	1
injury not obvious	3	2	2	1		2	
coffins	3	1		1	1	1	
death, see face with blood	2		3		1		
death, face obscured	1	1	2		1	1	
death, see face no blood	4		1				
death disembodied		2	1				
metaphorical	1						
Total	60	28	28	22	18	17	2

Table 2: Number of photographs depicting Iraqi civilian injury, grief and death. (Photographs depicting journalist or other non-Iraqi casualties have been excluded from this data.) (Parry, 2012)

The totals presented here are relatively small so only numbers rather than percentages are included. However, such data does allow us to track how and where depictions of injury and death appear across the sample of seven newspapers. The database also facilitates designing a query to find which photographs appeared on the front page; and whether the caption or headline attributed blame. A key finding here was that despite the global circulation of digital images, the differences regarding prominence of photographs and their graphic nature was determined not by technological or access issues but by editorial judgements: 'These are guided by the broad adherence to political positions on the war and generally conservative concerns for a perceived benchmark of 'decency' – concerns that are aligned with toning down criticism of 'our boys' engaged in military action' (Parry, 2012: 184).

There is not the space here to go into all the possible variations or analysis. The main point is that the numerical data provide the foundation to explore further how the newspapers depicted the human cost of war, and how each newspaper addressed its readers in terms of protecting their sensitivities, or of imagining their moral outrage.

From mass media to social media: what now counts as content?

The term 'mass' has become an outmoded term as media forms have proliferated especially with the arrival of digital technologies, and familiar patterns of media use are no longer characterized by sitting together to watch television or regularly reading the same print newspaper. And what we mean by 'media content' has also changed: it is not necessarily a recognizable genre produced by traditional media companies, with digital technologies enabling the cheap and easy creation and sharing of 'user-generated content'. YouTube videos can be devoured on the go, with the next 'recommended' video lined up and ready to play. I do not intend to explore the full implications of such socio-technological shifts here, but want to make the point that new media forms and genres, as well as changing audience habits, bring new challenges and opportunities for visual media content analysis.

Content analyses interested in the visual have long collated data 'beyond' the text. For example a content anlysis of film might include statistical data about distribution and international box office performance, in addition to character, plot and themes. Internet based communication technologies have not only made such information easily accessible for researchers, but also revolutionized the kinds of information we can extract directly from online platforms. For example, the video sharing website YouTube has opened up new possibilities for those who want to analyze not only the form of the video messages, but the 'metadata', such as the person who uploaded the video, or how many times it has been watched. Indeed, as with any web space designed for interaction, it is also often the

communicative purposes of the comments posted *below* the video that attract scholarly attention alongside the video content.

Broadcast Yourself: Analysing YouTube video content

Given the popularity and accessibility of YouTube, in terms of both uploading 'user' content and its tendency for short snappy videos viewable on the move, it has attracted attention from those interested in its educational and health communication potential. One such example is Yoo and Kim's (2012) study of obesity-related content on YouTube, which found that stereotypical depictions of eating behaviours persisted in this new media space, although they recognized that content analysis alone cannot tell us how the viewers engaged with such content. YouTube's original tagline was 'Broadcast Yourself' and its 'About' page now states, 'Our mission is to give everyone a voice and show them the world'. It is crucial then to think of YouTube not simply as a one-to-many communication channel, but as a space for multiple voices to share experiences and perform identities across geographical boundaries.

Content analyses designed with this in mind are those most adapted to the affordances of web platforms. In their analysis of YouTube videos uploaded by activists in response to the controversial anti-Islam film *Fitna*, van Zoonen et al. (2010) were able to collect relevant videos using an e-research tool that identified unique uploads (thus preventing double-counts) and automatically coded all metadata such as 'date of upload, gender, age, religion and country of origin as registered by the poster' (253). This inventory was combined with manual coding of 'length of the video, position of the video on Islam (positive, negative, unclear), position of the video on *Fitna* (positive, negative, not about *Fitna*, unclear), number of views, number of comments and the 'genre' of the video' (253). The authors were interested in how those posting videos were performing a form of citizenship; for instance, apologizing for *Fitna* as Dutch citizens or presenting an understanding of Islam as a peaceful

religion. Not only was this study alert to the possibility of using digital methods to extract data, it also explored the genres of the videos which were unique to digital culture '(tagging/jamming, cut-and-mix and vlogs)' (259), noting a sense of 'connectedness' among the video-posters and commenters across all styles of video address.

The challenges and opportunities of 'big data' and content analysis

Since the above-cited article was published we have seen the rapid global expansion of both social media usage and digital research methods. With so many users of social media platforms, conducting a visual content analysis of a certain issue or event could entail initially working with thousands or even millions of posts or tweets containing visual content. Traditional media research techniques have been designed with limited media outlets in mind, manageable via human coding alone. But the 'big data' generated by online media practices requires appropriate digital tools and methods, including automated coding or sorting. Web analytics technologies not only help the commercial sector track visitors to their websites, they are also one of the tools used by researchers to track engagement and interaction with certain content.

Due to the way in which software packages regularly appear and disappear, and with shifting 'terms of service' policies for popular platforms such as Twitter or Facebook restricting the way data can be gathered, I will not refer to certain programmes here. Instead there are research centres which specialize in digital methods, such as the Digital Methods Initiative at the University of Amsterdam or the Statistical Cybermetrics Research Group at the University of Wolverhampton, along with various resources collated by Deen Freelon of the University of North Carolina (http://dfreelon.org/). In some cases the data collection requires programming skills, whilst other software handily presents the tweets, comments or

images in an accessible format. Conducting computer-assisted content analysis brings its own challenges, not least when posted material disappears or is removed by moderators – content analysis requires consistency and the ephemeral nature of some web-based materials can hamper the requirements for validity. This means that one important extra step could be creating an archive of your sampled material to ensure it remains accessible and stable.

There are then practical and technological considerations to address in designing a computer-assisted content analysis. But thinking beyond practicalities, there is a deeper epistemological question about the nature of the content the analyst collects through web analytics tools and how this relates to the research objectives underpinning any study. Such studies might conduct content analysis of visual material in a radically different way to traditional approaches and with guiding questions which are focused more on understanding the sharing practices among 'networked publics', especially in times of crisis, conflict or political upheaval.

Visual social media content analysis: playing catch-up once again

Studies of visual social media have suffered from a familiar pattern, in the sense of lagging behind those studies focused on text. In part this is due to the nature of the data that has generally been accessible through the APIs (application programming interface) of platforms such as Twitter. As Highfield and Leaver (2017: 48) point out, '[t]he visual adds levels of trickiness to such analyses: first in accessing the images, videos, or other linked and embedded files, and then in studying them, which requires more individual intervention and interpretation than samples of 140-characters'.

Image-sharing has become central to the communicative practices on social media, whether of photographs, memes or animated GIFs (Graphics Interchange Format). This means that the focus for such studies has shifted: to genres of self-representation such as

profile pictures on Facebook (Emmons and Mocarski, 2014), or 'selfies' on Instagram (Veum and Undrum, 2017), but also to the ways in which more traditional forms of political campaigning such as posters have moved online (Lee and Campbell, 2016). In one sense there is continuity in the nature of questions driving the research – the ideological meanings in gender representations or the forms of visual symbolism designed to engage support for a cause – but there is also change, perhaps most clearly in the capture of interactivity (shares, likes, replies) which, even if in the most rudimentary way, integrates audience engagement into a 'content' study.

As mentioned, combining web analytics with human coding allows the researcher to record visual prominence in new ways. Whereas front page status and image size used to be staples for the content analysis design, once the affordances of social media are taken into account, the analyst instead counts the number of uploads, retweets, or likes, for example. The research tools might also establish the most 'influential' users, based on number of followers, and/or amount of content later shared by others. But reliability and validity concerns can arise when 'data' arrives readily packaged or through automated systems where the collection methods remain opaque.

So what does this all mean for quantitative visual content analysis?

Social media practices are varied and complex, and so methodological developments are required which take into account the ways in which the cultural significance of the images are constituted online, and the contexts for the production and dissemination. Visual content analysis can certainly play a part in this, but as Highfield and Leaver argue, this requires bridging the 'capabilities of the computational and the cultural analysis' (2017: 50) and recognizing how the particular affordances of each social media platform engender certain behaviours and practices. Computational tools allow the researcher to discover the most

frequently shared images relating to a significant event. But it might thereafter be more meaningful to trace the movement of a certain prominent image across time and the network of users, and the degree to which its meaning transforms in interplay with textual content, as well as how it becomes refashioned (as a cartoon or meme). For example, studies of iconic images, or the memes based on the original photograph, often start with a basic image search or content analysis in order to trace how that specific image is circulated and altered across digital media: the compositional elements are examined, but also the ways in which its cultural and symbolic significance is reconfigured and expanded through various online recontextualizations (Boudana et al., 2017). In other words, our approach to 'content' is required to adapt to the characteristics and practices associated with each medium and genre, as well as the epistemological shifts regarding how images become meaningful in the digital media environment.

A key concern for those examining the visual information flows online is the way in which images are recontextualized by various communities of users, who might inaccurately label such visual material. In some cases this is willful deception, but in other cases vivid imagery can spread as evidential support for a belief or cause on dubious grounds (that is not actually depicting what it is labelled as showing). So another challenge for content analysts of visual social media is the verification of images: How might a content analysis address source-verification and even the motivations for such image-based hoaxes?

CONCLUSION

In the studies mentioned above, quantitative content analysis of the visual is a significant part of the research design; but rarely the sole method. Whether employing framing, multimodal, discourse or semiotic analysis of content and form, not to mention contextual considerations relating to production, technologies, practices or audiences, the richest studies tend to benefit

from combining methods. The rigour of a well-designed content analysis allows researchers to make authoritative observations about patterns or trends and is particularly valuable for comparative studies. The emphasis in the latter sections of this chapter on digital images and sharing practices is not intended to underplay the importance of studying traditional media content. Global media companies arguably remain the principal producers of the cultural artefacts with the greatest social and cultural impact. As visual researchers continue to be intrigued by the assorted ways in which people, events and issues are represented across the mediascape, employing a method that helps us to capture what is comprised in such media is, at the very least, a solid foundation for further explorations.

REFERENCES

Bell, Philip (2001) 'Content analysis of visual images', in T. van Leeuwen and C. Jewitt (eds.), *Handbook of Visual Analysis*. London: Sage, pp 10-34.

Bell, Philip and Milic, Marko (2002) 'Goffman's Gender Advertisements revisited: combining content analysis with semiotic analysis', *Visual Communication*, 1: 203–222.

Berelson, Bernard (1952) *Content Analysis in Communication Research*. Illinois: Free Press.

Berger, Asa A. (1998) *Media Research Techniques,* 2nd edn. London: Sage.

Bock, Annekatrin, Isermann, Holger and Knieper, Thomas (2011) 'Quantitative Analysis of the Visual', in E Margolis and L Pauwels (eds.), *Sage Handbook of Visual Research Methods*. London: Sage, pp. 265-282.

Boudana, Sadrine, Frosh, Paul and Cohen, Akiba A. (2017) 'Reviving icons to death: when historic photographs become digital memes', *Media, Culture & Society*, 39: 1210-1230.

Emmons, Betsy and Mocarski, Richard (2014) 'She poses, he performs: A visual content analysis of male and female professional athlete Facebook profile photos', *Visual Communication Quarterly*, 21: 125-137.

Fahmy, Shahira (2004) 'Picturing Afghan Women: A content analysis of AP wire photographs during the Taliban regime and after the fall of the Taliban regime', *Gazette: The International Journal for Communication Studies* 66: 91-112.

Fahmy, Shahira (2005) 'US photojournalists' and photo editors' attitudes and perceptions: Visual coverage of 9/11 and the Afghan War', *Visual Communication Quarterly,* 12: 146-163.

Goffman, Erving (1979) *Gender Advertisements*. New York: Harper Torchbooks.

Hall, Stuart (1975) 'Introduction', in A. C. H. Smith with E. Immirzi and T. Blackwell (eds.), *Paper Voices: The Popular Press and Social Change*. London, Chatto & Windus. pp. 11-24.

Hallin, Daniel C. (1986) *The Uncensored War: The Media and Vietnam*. Oxford: Oxford University Press.

Hansen, Anders, Cottle, Simon, Negrine, Ralph and Newbold, Chris (1998) *Mass Communication Research Methods*. New York: New York University Press.

Highfield, Tim and Leaver, Tama (2016) 'Instagrammatics and digital methods: studying visual social media, from selfies and GIFs to memes and emoji', *Communication Research and Practice*, 2: 47-62.

Hughey, Matthew (2014) *The White Savior Film: Content, Critics, and Consumption*. Philadelphia: Temple University Press.

Jacobson, Susan (2012) 'Transcoding the News: An investigation into multimedia journalism published on nytimes.com 2000–2008', *New Media & Society*, 14: 867-885.

Kress, Gunther and Van Leeuwen, Theo (1996) *Reading Images: The Grammar of Visual Design*. London and New York, Routledge.

Krippendorff, Klaus (2004) *Content Analysis: An Introduction to its Methodology*. London: Sage.

Larsen, Peter (1991) 'Media contents: Textual analysis of fictional media content', in K. B. Jensen and N. W. Jankowski (eds.), *A Handbook of Qualitative Methodologies for Mass Communication Research*. London and New York: Routledge. pp. 121-134.

Lee, Benjamin and Campbell, Vincent (2016) 'Looking out or turning in? Organizational ramifications of online political posters on Facebook', *International Journal of Press/Politics*, 21: 313-337.

Lutz, Catherine A. and Collins, Jane L. (1993) *Reading National Geographic*. Chicago: Chicago University Press.

Neuendorf, Kimberley A. (2002) *The Content Analysis Guidebook*. London: Sage.

Parry, Katy (2012) 'The First 'Clean' War? Visually framing civilian casualties in the British press during the 2003 Iraq invasion', *Journal of War and Culture Studies.* 5: 173-188.

Riffe, Daniel, Lacy, Stephen, Fico, Frederick G. (2014) *Analyzing Media Messages: Using Quantitative Content Analysis in Research*, 3rd edn. London: Lawrence Erlbaum.

Rose, Gillian (2012) 'Content analysis: counting what you (think you) see', in *Visual Methodologies: An Introduction to Researching with Visual Materials*, 3rd edn. London: Sage. pp. 81-104.

Schwalbe, Carol (2013) 'Visually framing the invasion and occupation of Iraq in TIME, Newsweek, and U.S. News & World Report', *International Journal of Communication*, 7: 239–262.

Sontag, Susan (1979) *On Photography*. Harmondsworth: Penguin.

Sun, Zhen (2017) 'Exploiting femininity in a patriarchal postfeminist way: A visual content analysis of Macau's tourism ads', *International Journal of Communication*, 11: 2624–2646.

van Zoonen , Liesbet, Vis, Farida and Mihelj, Sabina (2010) 'Performing citizenship on YouTube: Activism, satire and online debate around the anti-Islam video Fitna', *Critical Discourse Studies*, 7: 249-262.

Vultee, Fred (2007) 'Dr. FDR & Baby War: The world through Chicago political cartoons before and after Pearl Harbor', *Visual Communication Quarterly*, 14: 158-175.

Yoo, Jina H. and Kim, Junghyun (2012) 'Obesity in the new media: A content analysis of obesity videos on YouTube', *Health Communication*, 27: 86-97.