NB! This is a pre-print of a chapter due to appear in an edited volume on Amsterdam University Press.

To cite the article, please use the following reference:


If you have questions or wish to send me feedback about the chapter, feel free to contact me at tuomo.hiippala@helsinki.fi.
A multimodal perspective on data visualization

Tuomo Hiippala
Department of Languages
Digital Geography Lab
Helsinki Institute of Sustainability Science
University of Helsinki, Finland

1 Introduction

The current interest in multimodality, or how multiple modes of communication cooperate and interact, has opened up new opportunities for theoretical reflection and practical application within several fields. Thus we find in linguistics increasingly widespread statements that language should be seen as occupying just one form of communication among many other, equally important expressive resources; visual communication begins to consider aspects of language; art history takes in the moving image; human-computer interaction design is extended to include tactile and gestural communication rather than just language; and so on.

Research on data visualization, however, has not yet fully benefitted from the interdisciplinary perspective that defines most of the current research on multimodality. Schwalbe (2015) observes that although visualizations “lie at the nexus of journalism, graphic design, visual communication, cognitive science and computer science ... much of the research is somewhat dated” (Schwalbe 2015, 438, my emphasis). Previous research has established principles for visualising information (Tufte 1983, 1997), mapped the repertoires of visual expression (Bertin 1981, 1983, 2001, Engelhardt 2002) and explored the perception and reception of data visualizations (Holsanova et al. 2009, Ware 2012, de Haan et al. 2018).

Another stream of research has recently called for attention to how data visualizations may attempt to influence or even deceive the viewers (Dick 2015, Kennedy et al. 2016) and identified inequalities in access to the kinds of literacies needed for making sense of data visualisations (D’Ignazio 2017). These contributions have provided a much-needed critical perspective to complement the design- and reception-oriented approaches introduced above. At the same time, however, attempts to describe the multimodality of data visualizations have been relatively few (Engbretnsen & Weber 2017, Bateman et al. 2017), although Ledin & Machin (2018) have argued that any form of critical inquiry that targets contemporary forms of communication, such as data visualizations, must now be supported by a comprehensive theory of multimodality. This argument resonates with the oft-cited quote from Halliday (1994), who observed that:
“A discourse analysis that is not based on grammar is not an analysis at all, but simply a running commentary on a text.” (Halliday 1994, xvi)

Although the notion of ‘grammar’ has been suggested as problematic for multimodal analysis, because it relies on strong assumptions about form, a property that visual modes of expression do not necessarily respect (Bateman 2014a, 168), Halliday’s observation about the need for a solid theoretical foundation for understanding the phenomenon under analysis remains crucial. For this reason, multimodal analyses are not only highly valuable on their own right for increasing our knowledge about multimodality as a phenomenon, but can also support critical perspectives to data visualizations by placing these analyses on a robust, multimodally-informed foundation.

Acknowledging the fundamental role of multimodality means subscribing to the view that data visualizations are more than simple combinations of pie charts, scatter plots and other graphs with decorative illustrations and typographically-emphasised text that may be simply read off some medium of presentation. A multimodal approach focuses on how meaningful combinations of written language, illustrations, photographs, diagrams, maps, layout and other modes of expression emerge in data visualizations. How such combinations are supported multimodally and become interpretable across a wide range of different media remains a wide and open research question. For data visualization, answering these questions undoubtely requires a theory capable of taking on the various media in which data visualizations appear.

In this chapter, I aim to show how the framework proposed in Bateman et al. (2017) can be used to lay a foundation for analysing the multimodality of data visualizations. The proposed approach starts with a detailed account of the medium of presentation, the modes of expression provided by the medium, and what kinds of engagement the combination of media and modes demands from those interacting with them. By doing so, I attempt to show how much detail a comprehensive theory of multimodality can pick out for discussion as a part of critical inquiry (cf. Ledin & Machin 2018). I do not, however, seek to pursue any critical analyses myself, but attempt to support this endeavour as outlined above.

2 Media and their canvases

Data visualizations are used in different kinds of communicative situations across a wide range of media. They are presented on websites, printed on newspapers, shared on social media feeds and projected on public displays, to name just a few examples (Lima et al. 2014, Boumegr et al. 2017, Amit-Danhi & Shifman 2018). For this reason, identifying the medium in which the data visualization is presented is a natural first step for their analysis, which has far-reaching consequences for a description of their multimodality. However, if media are characterised purely on the basis of their physical or technological characteristics, for instance, by setting up dichotomies such as ‘print’ or ‘digital’, we risk oversimplifying the medium in which data visualizations appear (see e.g. Elleström 2010, Bateman 2017).
In order to break down the abstract concept of ‘media’ (or medium) and prepare it for multimodal analysis, Bateman et al. (2017, 86–87) adopt the notion of a canvas to describe any potential carrier of semiotic modes that may be taken up for interpretation. The notion of a canvas places very few demands on the underlying materiality – almost anything capable of carrying intentionally-produced signs will do. Thus a note scribbled on a napkin is just as interpretable as a daily menu written on a chalkboard, because the presence of semiotic modes signals that the canvas in question is being offered up for interpretation. Multimodality research conceptualises semiotic modes as socially-shaped resources for making and exchanging meanings, and just like the semiotic modes, the canvases provided by a medium come to be by virtue of being embedded within a community of users (Bateman 2011, Kress 2014).

Bateman et al. (2017) propose that physical or technical media may be characterised as recognisable ‘bundles’ of canvases defined by patterns of production and consumption. To exemplify, the producers of audiovisual news broadcasts often allocate parts of the canvas to news tickers, stock and weather information, etc., in addition to the actual broadcast (Tan 2011). From the consumer’s perspective, identifying these canvases does not only generate expectations towards what kinds of communicative situations may take place on them, but also anticipates the semiotic modes most likely to be encountered in a particular communicative situation. In order to characterise the properties of a canvas, Bateman et al. (2017, 104) propose accounting for several material properties: space (2D or 3D), temporality (static or dynamic), transience (permanent or fleeting) and how the user is positioned with respect to the canvas (distanced observer or immersed participant).

Because some of these affordances are inherited from the materiality of the medium, this is also where differences begin to emerge between canvases. Nevertheless, all canvases that carry data visualizations must have an inherent spatial (2D) extent, which provides access to expressive resources provided by layout (Waller 2012). First differences emerge within the temporal extent: spatial canvases without temporal extent are considered static, whereas their counterparts with a temporal extent can be characterised as dynamic. Dynamic canvases may be either immutable or mutable, which also determines their degree of interactivity (Weber 2017, 246–247). In most cases, these canvases are also designed, which in this context implies that the content (or underlying data) cannot be altered by the user.

For the multimodal analyst, being able to pick out canvases and their properties for closer analysis is crucial for making sense of how the underlying medium is used to support a data visualization. This is, in fact, ultimately necessary for establishing differences between data visualizations presented on their own dedicated websites (see e.g. Zambrano & Engelhardt 2008, Bounegru et al. 2017) and those embedded in social media feeds, or for capturing the differences between information graphics in printed newspapers and their counterparts in digital media (see e.g. Lima et al. 2014). In both cases, the theoretical apparatus must be capable of taking on the complexity of the communicative situation in which a data visualization is mobilised, as opposed to hiding it from view. This is why the following section introduces an additional perspective,
which attempts to capture the kinds of interaction demanded by canvases, shifting the
attention from the production of canvases to their consumption.

3 Exploration and composition

Dynamic data visualizations and static information graphics have been suggested to de-
Bateman et al. (2017, 105) characterise this engagement in terms of ergodic work, re-
defining the concept originally developed by Aarseth (1997) for multimodality research.
The concept of ergodic work seeks to characterise a communicative situation in terms
of the effort required from those participating in the situation. More specifically, the
concept emphasises how participants co-construct the communicative situation they are
interacting with/in and to what extent the participants may manipulate the situation
(cf. Bucher & Niemann 2012). Because communicative situations can take place on can-
vases embedded within one another, different canvases may demand different forms of
ergodic work.

As a form of ergodic work, engaging with interactive data visualizations may be
broadly characterised as exploration (Bateman et al. 2017, 108). Exploration involves
substantial ergodic work on behalf of the viewer, in the form of interacting with the
visualization, for instance, by choosing which parts of the underlying data are rendered
by manipulating the data visualization via an interface. The extent to which the vi-
sualization may be manipulated is naturally determined by its degree of interactivity
(Weber 2017, 246–247). What remains beyond the user’s reach, however, is the underly-
ing data. In other words, the presentation of the data may be altered, but not the data
itself. For this reason, the communicative situation of engaging with an interactive data
visualisation may be characterised as ergodic yet immutable (Bateman et al. 2017, 108).

Another form of ergodic work required for interpreting data visualizations is that of
composition, which requires the viewer to determine how the information presented on a
2D canvas is to be put together. This involves selective visual perception and interpre-
tation, which may be revealed using methods such as eye-tracking, as Holsanova et al.
(2009) and de Haan et al. (2018) have shown for data visualizations in printed and digital
newspapers. It should be noted, however, that as forms of ergodic work, composition
and exploration are not mutually exclusive. In fact, exploring a data visualization must
necessarily involve ergodic work in the form of composition, as interpreting an interac-
tive data visualization involves making sense of information rendered on a 2D canvas at
a given point in time. These embedded forms of ergodic work emerge naturally from
canvases embedded within each other (Bateman et al. 2017, 109).

To summarize, the concept of ergodic work seeks to draw attention to how data
visualizations set up different types of communicative situations. As the following anal-
yses will show, many of these differences may be traced back to the properties of the
physical/technical medium, because the medium provides the canvases on which these
communicative situations take place.
4 Three example analyses

In this section, I demonstrate how the procedure set out in Bateman et al. (2017, 228) can be used to identify canvases in three different data visualizations, in order to lay a foundation for their multimodal analysis. All three examples discussed below fall under the umbrella of sustainability, such as biological conservation, global warming and marine pollution. The examples feature contributions from various semiotic modes in the form of written language, photography, diagrams and graphic elements. For current purposes, I do not seek to pursue a detailed analysis of their structure and functions, but characterise them rather broadly. The same applies to any discourse relations that hold between them. In contrast, by focusing on the canvases I seek to provide the means for increased analytical control, laying a foundation for more detailed analyses.

4.1 Static information graphics

Figure 1 shows an information graphic – here understood as a form of static data visualization – produced by Graphic News, a London-based agency that produces news graphics for media outlets around the world. The information graphic combines several modes of expression – written language, photography, maps, two-dimensional illustrations and diagrammatic elements – which are organised on several overlapping canvases. The wealth of semiotic modes present exemplifies why information graphics may be conceptualised as a composite semiotic mode, which provides the ‘glue’ necessary for joining together contributions from individual semiotic modes (Bateman et al. 2017, 289). This ‘glue’ may be traced back to a specific form of discourse semantics that support the formation and interpretation of composite units, which draw heavily on layout as a resource for expression (Bateman et al. 2017, 264).

For this reason, interpreting information graphics requires ergodic work in the form of composition, as described in Section 3. The viewer must identify the semiotic modes, consider their specific contributions and relate them to each other in the layout space. Resolving the discourse relations between written language and the photograph in the upper part of the graphic is a fairly trivial task, as written language is used to provide a headline, background information and to identify the rhino in question. Making sense of the lower part, in contrast, may prove more challenging due to discourse relations that hold between multiple semiotic modes presented on several overlapping canvases, which are a common feature of information graphics (Bateman et al. 2017, 291).

The lower part of the visualization features a map that shows the current and historic spatial distributions of rhino populations. MacEachren (2004, 317) notes that maps use overlays to present complex phenomena in space and time, but required processes of attribution – assigning meaning to the overlays – are often dependent on other modes of expression. This process of attribution is exemplified in Figure 1 by the accompanying legend, which uses coloured graphic elements and written language to group together different species of rhino and establishes their current and historic spatial distributions. Laid out on top of the map is another canvas, which provides additional information on specific rhino populations using combinations of two-dimensional illustrations, written
Figure 1: A static information graphic reporting on the death of the last male northern white rhino, produced by Graphic News.
language and diagrammatic elements, such as lines and containers. In addition, the diagrammatic mode is used to add information to the description in the upper part of the graphic by locating the Ol Pejeta Conservancy and the historic range of the northern white rhino.

These discourse relations, which are drawn between contributions from multiple semiotic modes and extend across the canvases, could be described in detail using various multimodal frameworks developed for this purpose. This level of description, however, is beyond the scope of this chapter (for a recent overview of this area, see Bateman 2014b). In order to prepare for drawing comparisons between static and dynamic 2D canvases, it is worth noting how the static information graphic negotiates the limited layout space by using overlapping canvases. As the following examples will show, this limitation is largely absent from dynamic data visualizations, which can exploit material properties such as temporality and transience to increase the available layout space.

4.2 Non-interactive dynamic data visualizations

Figure 2 features four screenshots captured from a dynamic data visualization, which illustrates temperature anomalies by country between 1900 and 2016. Created by Antti Lipponen, a researcher at the Finnish Meteorological Institute, this 2D visualization is centered around a circular structure with the data – represented by coloured bars – and their respective labels laid out on concentric circles. The labels (countries) and bars (observations) are organized along the concentric circles according to geographical location. In addition, the top right-hand side of the visualization features a line graph which shows the global average for each year, summarising the individual observations presented using the circular bar plot.

The visualization, which uses two semiotic resources provided by the diagrammatic mode to represent time series data, namely circular bar plots and line graphs (Tversky 2017, 350). The visualization may be divided into three distinct canvases, which differ in terms of their temporality and transience. The first canvas, which carries the header, data sources and author information, all positioned along the edges of the visualization, is static and permanent. The second canvas in the middle of the visualization is dynamic and fleeting in terms of temporality and transience, which enables the circular bar plot to be rendered again at each time step. Finally, the third canvas on the top right-hand side is also dynamic but permanent, which allows the line graph to be updated at each time step.

This difference in transience may be traced back to the diagrammatic resources and the kind of communicative work they are intended to do. Whereas the circular bar plot is used to represent changes among a large number of simultaneous observations, the line graph tracks a single observation over time to summarise the trend. The permanent canvas allows the line graph to use the two-dimensional layout space to keep all previous observations in view, which is something the circular bar plot cannot do: rendering each time step on the circular bar plot is simply not feasible due to limited layout space, the obvious solution is to distribute the representation over time, which is enabled by the fleeting canvas. Despite rapid changes, tracking changes on this canvas is facilitated by
the way the human brain prioritises the processing of colour and line length (Ware 2012, 154–155).

In terms of ergodic work, this visualization may be characterised as a dynamic data visualization, whose interpretation does not entail exploration, but constant composition. Exploration is not required, as animated graphics are not interactive, and consequently cannot support manipulation or navigation by the user (Weber 2017, 247). Depending on whether the visualization is opened in a web browser or a media player application, initiating the temporal sequence may involve clicking a play button, but this interaction emerges from the medium in which the visualization is realized, not the visualization itself. Such low-level interactions are commonly used for imposing control over embedded dynamic canvases in digital media, and should not be confused with interactivity inherent to the data visualization (cf. Hiippala 2017, 424–425), which are taken up for discussion below.
4.3 Interactive dynamic data visualizations

The final example in Figure 3, *The Seas of Plastic*, is an interactive dynamic data visualization created by *Dumpark*, a design agency based in Wellington, New Zealand, which visualises how plastic pollution accumulates into large circulating gyres in oceans. To do so, the visualization provides two distinct views – designated as map and source views, respectively – which are both presented on their own canvases. These canvases may be rendered for viewing via the interface in the top-right corner of the embedding canvas, which remains constantly visible to the user.

In addition to the interface for exploring the visualization, additional levels of interactivity are introduced on the two canvases. The map view, shown in the upper part of Figure 3, features a 2D representation of a globe that may be freely rotated by clicking and dragging. A legend, positioned in the lower left-hand corner, is used to attribute meaning to the overlays rendered on the globe, which bears close resemblance to the discourse relations in the information graphic in Figure 1. The user may also select a specific gyre on the right-hand side interface, which rotates the globe into a position that shows the selected gyre. Selecting a source or a gyre in the source view highlights coloured bands that show the source or destination of plastic pollution. Multiple sources or gyres may also be selected simultaneously for drawing comparisons between them.

Together, multiple user interfaces on several canvases lend this data visualization a high degree of interactivity. According to Weber (2017, 247), this entails that the users are allowed to explore the visualization, interact with the data and influence its representation, which corresponds closely with what Bateman et al. (2017, 108) characterise as ergodic work in the form of exploration (see Section 3). At the same time, the contributions from various semiotic modes and the discourse relations that hold between them closely resemble those found in static information graphics and non-interactive dynamic data visualizations: written language provides additional information on graphics, legends accompany cartographic representations, colour creates distinctions, etc., which must be decomposed and put back together for interpretation. In other words, making sense of the interactive data visualization requires ergodic work both in the form of exploration and composition, a feature which separates it from static information graphics and non-dynamic data visualizations.

5 The need for exhaustive analyses

Engebretsen & Weber (2017, 289) have recently argued that a multimodal account of data visualization must move beyond identifying which semiotic modes are used and for what purpose, and move towards a closer analysis of production and consumption, in order to pinpoint how meaning potential emerges. This argument is very similar to what Bateman et al. (2017, 221–222) propose for multimodality research in general, underlining the need to pursue analyses in an exhaustive manner, which Ledin & Machin (2018) identify as a key component of any critical inquiry as well. Such analyses should involve accounting for the communicative situations in which the visualizations partic-
Figure 3: The Seas of Plastic, an interactive dynamic data visualization by dumpark.com
ipate, identifying and describing the canvases on which these communicative situations take place, the semiotic modes these canvases provide, the genres that determine how the semiotic modes are structure, etc. This does not, however, necessarily entail full-blown analyses at each stage, but can also serve as a tool for limiting the scope of investigation.

Figure 4: The decomposition of (1) static information graphics, (2) non-interactive data visualizations and (3) interactive data visualizations into canvases.

That being said, identifying the canvases and describing their properties can be proposed as a first step towards a more comprehensive analysis of production processes. As Section 2 showed, canvases inherit affordances from the materiality of the medium that carries them, and they may be manipulated in different ways for different communicative purposes. What motivates the producers to manipulate these canvases and their material affordances can be revealed using ethnographic methods (Hiippala 2016, Zha 2017). For the examples discussed above, the properties of the canvases are visualised in Figure 4, in which they are marked as being either present (+) or absent (-).

The static information graphic (1) in Figure 4 illustrates how certain material properties of the medium are passed down to all canvases. The consequences are clear: a 2D medium without a material prerequisite for a temporal dimension can never be used to instantiate a dynamic canvas. This consequently rules out any semiotic modes that require this property. In contrast, the digital medium in data visualizations (2) and (3) are realised affords controlling temporality and transience of any canvas, which also provides the foundation for interactivity. To summarise, the canvases and their properties determine which semiotic modes may appear on them, and thus their description should
precede any in-depth description of the semiotic modes used and their contribution to the visualization at hand.

One contribution that emerges from mapping the canvases at play is the role of layout. Bateman (2008, 157) proposes the term page-flow for describing the semiotic mode responsible for setting up discourse relations in the layout space, which hold between contributions from distinct semiotic modes. The discourse semantics of page-flow are exemplified, for instance, by the relations that hold between the diagrammatic overlay and the underlying map in the static information graphic in Figure 4. The role of page-flow in organising the spatial structure of 2D canvases is highlighted by indicating page-flow as the active semiotic mode in all visualizations in Figure 4. However, to what extent the discourse semantics of page-flow differ between data visualizations and entire page-based documents remains an open question.

To sum up, the widespread use of information graphics and data visualizations makes their multimodal analysis challenging, given the wealth of communicative situations in which they appear. However, only a sufficiently developed theoretical apparatus, which is able to impose control on the communicative situation in which the visualizations appear, can advance our understanding of how the visualizations work. This will undoubtedly require an extensive programme of empirical research, which must involve specialists from various fields, given the need for exhaustive analyses that cover the whole range of phenomena from production to consumption (cf. Waller 2012, Zha 2017).

6 Conclusion

In this chapter, I have attempted to highlight how much state-of-the-art theories of multimodality can reveal about data visualisations even before venturing into in-depth descriptions of semiotic modes and the discourse relations that hold between their individual contributions to the data visualization under analysis. By drawing on the notion of canvas, recently introduced in Bateman et al. (2017), I have also sought to establish a foundation for further analyses by attending closely to the underlying properties of the medium, investigating their contribution to meaning-making, as called for by Ledin & Machin (2018). Ideally, such multimodally-informed insights should provide a basis for critical insights into the use of data visualisations in society, allowing them to be strongly rooted in well-informed analyses of multimodal discourse. Supporting these critical perspectives will also require continuous refinement of multimodal theories that are applicable to data visualisations, and given the rapid spread of data visualisations into all areas of society, these theories must undoubtedly be founded on empirical research.

References


Zambrano, R. N. & Engelhardt, Y. (2008), Diagrams for the masses: Raising public