3. Inventorizing, situating, transforming: Social semiotics and data visualization

Giorgia Aiello

Abstract
This chapter is an overview of social semiotics as a productive framework for research on data visualization. It provides conceptual instruments that can be used to explore the relationship between the formal properties of data visualization and the meanings and practices that these may promote or hinder among users. In particular, the chapter argues that a social semiotic framework can be used to inventorize, situate, and transform visualization resources. Overall, it links descriptive, interpretive, and critical objectives to generate a framework aimed at understanding how data visualization ‘works’ from a formal standpoint, what meanings are consistently associated with particular semiotic resources, and how both key semiotic ‘rules’ and dominant meanings may be questioned and changed.

Keywords: Social semiotics; Data visualization; Semiotic resources; Visualization design

Introduction
This chapter is a focused critical overview of social semiotics as a productive framework for research on data visualization. It aims to provide conceptual instruments that can be used to explore the relationship between the formal properties of data visualization and the kinds of responses, engagements, and practices that these may promote or hinder among users. Over the last decade or so, and in the wake of digitalization and datafication, data visualization has emerged rapidly as what Engebretsen and Weber (2017) have defined as a ‘super-genre’ that is used to accomplish a wide variety of communicative tasks across an increasing number of professional and
institutional communities of users. Beer and Burrows (2013) highlight that, as a whole, we have witnessed the rise of a ‘visualization of culture’ and a ‘culture of visualization’ across spheres of social activity and cultural production, as ‘there are not many things that have yet to be visualized and archived’ (p. 62). Different kinds of data visualization have become privileged signs to mark the rationality of particular processes and promote specific attitudes towards various aspects of everyday life, ranging from policymaking to personal productivity. As Ledin and Machin (2018) point out, often diagrams, charts, and other types of visualization are used not only to illustrate how things are but also, above all, ‘how things should be done’ (p. 335).

Precisely because of the increasing social significance of this phenomenon, there is a growing body of academic literature centred on critical, practical, and combined approaches to the formal and overall aesthetic qualities of data visualization. Generally speaking, these approaches offer very useful insights to examine data visualization design from an ideological, professional, or praxis-based standpoint. On the one hand, it has become increasingly urgent to examine what Kennedy and Hill (2017) define as the ‘visual sensibilities’ (p. 2) that are at work in the ways in which ordinary people respond culturally and engage emotionally with data and their visualizations. On the other hand, professional and institutional uses of data visualization techniques must be examined in the light of their underlying histories, conventions, and changes over time and across contexts. For these reasons, a detailed appraisal of data visualization’s main semiotic resources, or its tools for meaning-making, is key to empirical research in this field. Unlike other currently more widespread approaches to data visualization research rooted in cultural and social theory, a social semiotic approach focuses keenly on the formal properties of visualizations together with their semiotic and social affordances.

As I will explain later, a social semiotic approach entails a systematic mapping of semiotic resources together with an empirical if not ethnographic investigation of how such resources came to be the way they are, how they are used or understood by a variety of individuals and groups of people, and how they are shaped by dominant practices and regulated by given institutions. It is in this sense that social semiotics is inherently critical, as it relates texts to contexts to reflect on the social and political implications of meaning-making. However, social semioticians are concerned not only with the politics but also with the potentials of semiosis. Ultimately, one of the major aims of social semiotics is to contribute to semiotic innovation, or envision ways in which the ‘rules’ of sign-making may be broken or
changed (van Leeuwen, 2005). This matters because semiotic innovation can contribute to engendering social change. In this chapter, I therefore argue that a social semiotic framework of this kind can and ought to be extended further to *inventorize, situate, and transform* the semiotic resources associated with data visualization.

To explore relevant conceptual tools that are central to social semiotics as a mode of inquiry, then, the chapter begins with a broad discussion of the methodological dimensions of social semiotics, together with an initial discussion of existing scholarship in this area. I then delve into three main theoretical and analytical areas. First, I outline some of the major sources and methods that we can harness to begin inventorizing data visualization resources. In doing so, I review a selection of analyses of relevant multimodal semiotic artefacts and technologies such as diagrams (Ledin & Machin, 2016a and 2016b; Bateman et al., 2017), infographics (Bateman et al., 2017; Amit-Danhi & Shifman, 2018), and PowerPoint (Djonov & van Leeuwen, 2013; Zhao et al., 2014). Second, I explain how data visualization resources can be situated in their contexts, particularly through historical and ethnographic approaches. Finally, I advance the idea that social semiotics can contribute to transforming data visualization resources. The overall aim here is to link descriptive, interpretive, and critical objectives to generate a framework aimed at understanding how data visualization ‘works’ from a formal standpoint, what meanings are consistently associated with particular semiotic resources, and how both key semiotic ‘rules’ and dominant meanings may be questioned if not changed.

**Why social semiotics?**

Critiques of data visualization often focus on the truth-making claims and related epistemological implications of its design (see Halpern, 2015). For example, recently Gray et al. (2016) explored some of the ways in which data visualization’s ‘ways of seeing’ and ‘ways of knowing’ can be understood in relation to ‘the aesthetics, cultures, values, ideals and practices associated with their production’ (p. 294). When it comes to research on the visual and multimodal detail of data visualization, there is still a predominance of practice-based research. Edward Tufte’s groundbreaking work on the design norms underlying the visual display of information has been both widely criticized and surpassed by technological and cultural changes in how visualizations are both produced and used (see Tufte, 1983, 1997). Levels of interest in research on the ‘good practices’ of data visualization design have
grown among praxis-oriented thinkers. In his data visualization handbook, for example, Andy Kirk (2016) offers guidance on the development of design solutions across the ‘five layers of the visualisation design anatomy’ (p. 145), which he defines as data representation, interactivity, annotation, colour, and composition.

Synthesizing critical and practice-based approaches, Catherine D’Ignazio and Lauren Klein (2016) have claimed that theories from the humanities can be used to inform and change visualization design. Their contribution is largely focused on ensuring that the design process is inclusive and pluralistic at all stages, from the selection of data sources and representational strategies to the ways in which design teams are composed and the insights and experiences of end users are taken into account. And because they speak as part of a science and technology studies debate on dominant epistemological perspectives and power relations in data visualization design, D’Ignazio and Klein also primarily focus on structure and practice rather than form and meaning.

In the collaborative study on visualization conventions that I conducted with a group of researchers led by Helen Kennedy, we laid the foundations for a social semiotic approach to research on data visualization, with the explicit aim to understand how power works through some of the key semiotic resources found across visualizations (Kennedy et al., 2016). Likewise, Ledin and Machin (2018) propose a general framework for the study of ‘data presentation’ as a semiotic material, or a particular form of communication set apart by unique affordances and canons of use. In turn, Engebretsen and Weber (2017) highlight that data visualization is multimodal, as it is usually enacted as a deployment of multiple graphic modes including, for example, ‘typography, layout, maps, diagrams, and drawings’ (p. 279) together with colour as ‘an integrated component in all the other ones’ (p. 279). As they explain, in digital media data visualizations ‘can be static and monologic, but they can also be dynamic and dialogic’ (p. 289), they can be more or less explorative or open to interpretation, and they can be both pictorial or non-pictorial, with building blocks like photographs, illustrations, geometric shapes, and abstract motifs being equally available to visualization designers. This is important work, but nonetheless, there is little systematic research that combines both how data visualization design works semiotically and the politics and potentials of this semiotic work in relation to specific contexts and for particular groups of people.

As a methodology that is highly akin to critical discourse analysis, social semiotics is interested in what Caldas-Coulthard and van Leeuwen (2003) define as ‘the processes and products of discourse’ (p. 3), or both sign-making practices and their concrete outcomes together with their underlying ‘ways
of knowing’ and implications for our ‘ways of seeing’. In this sense, social semiotics is not merely a method or collection of methods, but rather a theoretical approach to empirical research. Like critical discourse analysis, and through a Foucauldian lens, social semiotics considers language and sign-making more broadly as key to the reproduction or transformation of social structures. However, social semiotics is also interested in how language and other modes of communication, particularly visuals, work together to make meaning.

Social semiotics originates from a synthesis of structuralist semiotics and Halliday’s (1978, 1985) systemic functional linguistics. Social semiotics is functionalist in that it considers all sign-making as having been developed to perform specific actions, or semiotic work (Hodge & Kress, 1988). Just like semiotics, it is also concerned with the internal structures of texts and, increasingly, also of other semiotic artefacts (e.g. architecture) and semiotic technologies (e.g. PowerPoint). Unlike traditional semiotics as well as other textual methodologies, social semiotics places emphasis on ‘how people make signs in the context of interpersonal and institutional power relations to achieve specific aims’ (MODE, 2012). In doing so, social semiotics therefore posits that the physiological and technological means (e.g. sound or imagery) that we use to communicate are to be examined as semiotic resources which can be, and in fact most often are, actively mobilized to achieve political, economic, and ideological ends.

This dynamic approach to defining key concepts extends to the notion of meaning, which is not fixed, and where semiotic resources ‘have a meaning potential, based on their past uses, and a set of affordances based on their possible uses’ (van Leeuwen, 2005, p. 285). The nature of such meaning potentials depends on concrete uses of semiotic resources in specific social contexts where their uses are governed by what van Leeuwen (2005) calls ‘semiotic regimes’. In other words, sign-making is regulated through social practices and guided by authority, expertise, or simple conformity in particular contexts. Hence, social semiotics is also able to account both for top-down power and bottom-up polysemy in relation to the uses of semiotic resources.

As I mentioned in the introduction, then, the critical aims of social semiotics are inherent in its approach to examining sign-making, which is always both descriptive and interpretive. Combining a systematic appraisal of semiotic repertoires with an understanding of how their meaning potentials are established over time and in context enables the analyst to understand how semiotic resources are shaped by power relations and, in turn, also ‘who made the rules and how and why they might be changed’ (Jewitt & Oyama, 2001, p. 135). In his book-length introduction to social
semiotics, Theo van Leeuwen (2005) explains that (social) semioticians do three main things. First, they ‘collect, document and systematically catalogue semiotic resources—including their history’ (p. 3). Second, social semioticians ‘investigate how these resources are used in specific historical, cultural and institutional contexts, and how people talk about them in these contexts—plan them, teach them, justify them, critique them, etc.’ (p. 3). Finally, they also ‘contribute to the discovery and development of new semiotic resources and new uses of existing semiotic resources’ (p. 3).

To this three-pronged definition, I would also add that social semiotics extends Roland Barthes’s original, though unfinished agenda in *Mythologies*, where he emphasized the need to create ‘an appropriate method of detailed analysis’ (Barthes, 1972, p. 9) to reveal and undermine the meanings established and perpetuated by the bourgeoisie, which he defined as ‘the essential enemy’ (p. 9). While Barthes’s definition of power and the status quo was specific to his time and intellectual background, social semiotics can still be seen as a way to carry out Barthes’s *semioclasm*, or a radical attack on the naturalization of signs followed by a more democratic redefinition of what widely shared semiotic practices may look like (Aiello, 2006).

With its ability to link texts with contexts, semiotic production with social action, and meaning with power, social semiotics is an especially congenial framework for research on data visualization. For these reasons, here I propose that a social semiotic framework should be used systematically to *inventorize*, *situate*, and finally also *transform* the semiotic resources of data visualization as a multimodal ‘super-genre’ in its own right.

**Inventorizing data visualization resources**

As a first step in our social semiotic approach, we must therefore begin by *inventorizing* the semiotic resources that are typical of data visualization across media and contexts. As van Leeuwen (2005) explains, ‘[t]o make an inventory we first need a collection’ (p. 6). In other words, we must identify and catalogue resources that are representative of data visualization as a whole. This is a particularly challenging task, both because uses of data visualization cut across a vast range of social spheres, and because the existing empirical base to systematically describe key data visualization resources is still thin.

To begin building an inventory of data visualization resources and their possible combinations, we can draw from existing social semiotic and multimodal studies of data visualization and of related semiotic objects. In the study led by Helen Kennedy mentioned earlier, we identify four key data
visualization conventions, namely two-dimensional viewpoints, clean layouts, geometric shapes and lines, and the inclusion of data sources (Kennedy et al., 2016). By the same token, in their recent book on visual analysis, Ledin and Machin (2018) examine different types of ‘data presentation’ through a social semiotic lens, including lists, bullet points, line graphs, bar charts, and flow charts. In this analysis, they identify a set of semiotic resources, namely paradigms, spatialization, vertical and horizontal orientation, graphic shapes and icons, temporality, and causality. Similar analyses of related semiotic objects like diagrams, infographics, and PowerPoint can also be useful in building an inventory of data visualization resources. This is not only because some of these are used in data visualization (e.g. diagrams) or are, at times, confused with data visualizations (e.g. infographics), but also because these analyses offer a discussion of findings and concepts that are useful for a social semiotic analysis of data visualization. What diagrams, infographics and PowerPoint have in common with data visualization is that they are all often used to relay ‘hard’ facts and key strategic points, usually with the aim to maximize an organization’s outputs and increase its competitiveness.

Research on diagrams has focused both on the features of diagrams as semiotic objects in their own right (Ledin & Machin, 2016a) and on the existence of a ‘diagrammatic mode’, which can manifest itself both independently (e.g. through charts, graphs, and schematic drawings, or ‘self-standing’ diagrams) but also in combination with other semiotic modes. Bateman et al. (2017) explain that the diagrammatic mode can work together with other modes so as to ‘form composite units’ (p. 279) that are often set apart by the ‘stacking’ of elements such as labels and connecting lines over illustrations, maps, or photographs. They argue that information graphics are the resulting ‘composite’ mode, as these provide the ‘glue’ to the ‘rhetorical relations between contributions from an equally wide range of semiotic modes’ (p. 294). In providing this rhetorical cohesion, information graphics rely not only on diagrammatic elements, but also and perhaps most importantly on layout space as a semiotic resource in its own right. Amit-Danhi and Shifman (2018) highlight that the composite nature of digital infographics is also increasingly mobilized to ‘embed a rhetoric of participation’ (p. 15), for example by letting users choose layouts and selections of data.

Along the same lines, Theo van Leeuwen’s collaborative work on PowerPoint highlights the increasing importance of semiotic resources that are typical of visual design, rather than traditional media as such, in everyday communication—such as typography, layout, colour, and texture (Djonov & van Leeuwen, 2013). In doing so, it focuses on inventorizing the resources that the software itself makes available by design, for example by privileging certain resources
and uses over others in its interface or help menu. This work contributes an understanding of the relationship between software and their uses, thus moving away from the notion of ‘text’ to investigate the relationship between semiotic technologies and semiotic practices (Zhao, Djonov, & van Leeuwen, 2014).

In addition to findings from existing analyses, practice-oriented publications like Andy Kirk’s data visualization design handbook or Alberto Cairo’s guide to information graphics and data visualization can offer a good starting point for the development of an inventory of the modes and resources that are used by designers themselves for the creation of ‘good’ visualizations (Kirk, 2016; Cairo, 2013). Finally, it is foremost through extensive empirical data collection both from a variety of media (e.g. news media, school textbooks, government websites) and in the field (i.e. through contact with designers, media professionals, and ordinary users) that we can build a systematic inventory of data visualization resources.

This first step of the social semiotic approach may be interpreted as an attempt to outline a ‘grammar’ of data visualization design, or what Machin (2007) defines as a ‘lexicon of elements that can be chosen to create meaning in combinations’ and ‘a finite system of rules’ (p. 185) for their combination. However, it would be problematic to think of such an inventory as a grammar, in that our goal here is not so much to understand how data visualization is and ought to be done, but rather what its major resources are, and how these are mobilized in particular contexts and for specific purposes (see Engebretsen & Weber, 2017).

Situating data visualization resources

Precisely for this reason, the next step of our social semiotic framework entails an attempt to situate data visualization resources in their social and cultural contexts. As Jewitt et al. (2016) explain, one of the main aims of social semiotics is ‘to understand the social dimensions of meaning, its production, interpretation and circulation, and its implications’ (p. 58). Both historical and ethnographic methods are often invoked as key to a social semiotic understanding of meaning-making. Cultural and social histories of a variety of resources—like, for example, colour—are used productively to locate their origins, understand the material, cultural, and political forces that shaped them, and trace their changes over time (see, for example, the history of the colour blue by Michel Pastoureau, 2001). However, fieldwork, and ethnographic research in particular, has often remained an ideal among social semioticians. One exception is my own work, in which
I have adopted a multi-sited ethnographic approach to investigate the practices, motivations, and outputs of image-makers like photographers and graphic designers (Aiello, 2012a, 2012b). As Marcus (1995) writes, when the object of ethnographic investigation is in ‘the realm of discourse and modes of thought, then the circulation of signs, symbols, and metaphors guides the design of ethnography’ (p. 108). Because of this focus on the social lives of signs, rather than of particular sites or communities, a social semiotic approach will entail a focus on data visualization as it is produced and used across different social and geographical locales.

This said, there is also much to be learned from existing and ongoing ethnographic studies of particular sites and settings in which data visualization is produced, used, or consumed. Alongside Helen Kennedy’s collaborative work on designers’ intentions and ordinary people’s responses with regard to data visualization, there is also a growing body of work on the production and uses of data visualizations in newsrooms (see Engebretsen et al., 2018). In this regard, a social semiotic approach to data visualization can also benefit from sociological research on digital and data journalism, in that it offers detailed accounts of the material resources, skills, and tools that are available to those who make decisions about data visualizations across news media (Fink & Anderson, 2015). This said, when interviewing participants, it is important that researchers ask questions not so much about the intentions, motivations, feelings, and overall actions of participants in relation to data visualization, but more specifically about how they use or interpret particular semiotic resources. This can be done through elicitation or reconstructive methods, where participants are asked to comment on particular texts (in this case, specific visualizations) that the researcher shares with them or asks them to share during the interview. Ultimately, asking questions about ‘the set of semiotic choices that typify a given context’ (van Leeuwen, 2005, p. 14) contributes both to understanding the context itself and the reasons why specific semiotic resources come to be the way they are. In situating visualization resources in their contexts, particularly through ethnographic fieldwork, researchers will often also come across ‘new’ resources, which will thus go to enrich and extend their initial inventory.

Transforming data visualization resources

The knowledge generated through the descriptive and interpretive stages of the social semiotic approach to data visualization leads to an understanding of visualization resources as part of broader cultural processes and
power relations. A third and final stage in this framework focuses both on the politics and potentials of data visualization. Major semiotic resources and their combinations can be transformed to break away from dominant ‘visual sensibilities’ and therefore also promote particular forms of social action and social change. As I highlighted earlier in the chapter, the goal of social semiotics is to interrogate as well as redefine sign-making. This is not considered to be a neutral process, but rather as having both power-laden origins and powerful implications.

It can therefore be useful to combine both critical and creative ends to understand how data visualization may be both part of what Fairclough (1995) has termed the ‘technologization of discourse’ and what van Leeuwen (2008) more recently defined as ‘the new writing’, or the new dominant language of multimodal communication. On the one hand, data visualization may be seen as part of a powerful impetus towards the standardization of semiotic resources for ‘the engineering of social change’ (Fairclough, 1995, p. 3). In other words, broader shifts in discursive practices are often aimed at changing the ways in which given institutions—e.g. news media, universities, and governments—and publics think and act in relation to particular issues. For example, Fairclough (1992, 1996) focused extensively on how language was used to promote and normalize both marketization and managerialism in public institutions like schools, universities, and hospitals. Through an analysis of how data visualization resources may be increasingly codified within and across institutions, and how such processes of semiotic codification may be tied to broader structures of power, we can begin to provide an evidence-based, sustained critique of the politics of data visualization. In this regard, for example, Ledin and Machin (2016a, 2016b, 2018) are currently building a body of work on how the discourses of performance management and marketized steering are recontextualized into increasingly ubiquitous ‘strategic diagrams’. These are used to translate values like competitiveness and accountability ‘into graphic shapes’ with ‘a clear logic of cause and effect’ (Ledin & Machin, 2016a, p. 323).

On the other hand, data visualization ought to be approached as evolving, rather than fixed or unchangeable. According to van Leeuwen (2008), in ‘the new writing’ the distinction between different semiotic modes is increasingly blurred and, in fact, their relationships are always expressed visually—for example, through layout and ‘cohesive uses of colour, typography and other stylistic elements’ (p. 132). Across types of media (e.g. websites, newspapers and magazines, institutional documents, and PowerPoint presentations), imagery now tends to be actively combined with writing and other semiotic resources. Hence, writing or images alone are no longer the most
Inventorizing, situating, transforming authoritative sources of information and persuasion in isolation from one another. Unlike the ‘old writing’, then, this new ‘language’ is grounded in principles of visual design (rather than image or word alone) that used to be relegated to professional niches such as web and graphic design. What this means is that data visualization is part and parcel of much broader semiotic practices that are increasingly shaped by normative discourses found in style manuals and formal teaching in art and design schools, but that are also learned through ‘best practice’ (or approaches that are widely accepted and prescribed as being most effective and sound) and built into semiotic technologies like office software.

These normative discourses regulate the uses of particular semiotic resources and users’ competencies in spite of ‘all-too-easy affirmations of boundless choice and endless creative opportunity’ (van Leeuwen, 2008, p. 135). This said, van Leeuwen (2008) also exhorts students and scholars of visual communication to investigate how these new ‘languages’ work in practice, to understand what they can and cannot do, and assess how homogenous or varied their applications and uses are in different contexts. For example, in addition to outlining guidelines for ‘good’ data visualization design, data visualization designers and their students can use a social semiotic approach to examine the histories of particular semiotic resources (e.g. colour, but also shape or layout) as well as understand how these may be used in different social and cultural contexts. Likewise, praxis-oriented scholars of data visualization may want to shift their attention from the broader power structures and work practices that shape data visualization design to include considerations about the ways in which key semiotic resources are used and interpreted by specific groups of people. In both cases, a social semiotic approach may offer an enriched outlook on how data visualization design ‘works’ in society—thus yielding practical insights into how to adjust and indeed also transform key formal characteristics for purposes like inclusion and equality.

Conclusion

Research on data visualization in society can benefit greatly from approaches that examine the formal—that is, both visual and multimodal—characteristics of visualization design in relation to their implications for how we ‘make sense’ of the knowledges, facts, and perspectives communicated by data visualizations. A social semiotic framework contributes to a systematic investigation of semiotic resources like colour and layout, for example,
together with visualization elements like graphs and charts. However, this is an approach that doesn’t stop at a description—however systematic and comprehensive—of form. Instead, it links form to context to understand how semiotic resources work in practice, what they mean and do in everyday life, and ultimately also how they might be changed to do good, or at least do better. By inventorizing and situating data visualization resources, we can build evidence aimed at engaging with the politics and potentials of increasingly dominant, transversal uses of data visualization. In this way, we can also contribute to transforming a range of semiotic practices related to the production and uses of data visualization in everyday life.

References


**About the author**

**Giorgia Aiello** is Associate Professor in the School of Media and Communication at the University of Leeds, United Kingdom. Giorgia’s research focuses on the politics and potentials of visuality, materiality, and multimodality. Her book *Visual Communication: Understanding Images in Media Culture* (with Katy Parry) was published by SAGE at the end of 2019.