What is at stake in data visualization? 
A feminist critique of the rhetorical power of data visualizations in the media 

Rosemary Lucy Hill 

Abstract 
Data visualizations are powerful semiotic resources, which, it is sometimes claimed, have the power to change the world. This chapter argues that to understand this power we need to consider the uses to which visualizations have been put. Using visualizations relating to abortion as a case study alongside Klein and D'Ignazio's notion of a 'Bring Back the Bodies' in data visualization, I argue that visualizations tell a narrow story, removing contextual detail and omitting to ask questions important to women's health. To grasp the significance of this I propose a new body issue: the neglect of the viewer and those affected by decisions taken based on visualized data. Far from being a simple device to graphically display numerical data, therefore, there are important social and ethical issues at stake in data visualization.

Keywords: Abortion; Data visualization; Feminism; Bodies

Introduction

What is data visualization for? Data visualizations in the media are not just about giving people easy or pretty access to information. They are about telling stories and they therefore work within the narrative frames of their designers and disseminators. When influential data visualizers write that data visualization can 'change the world' (Kosara, Cohen, Cukier, & Wattenberg, 2009), implicitly for the better, we therefore need to ask questions...
of what they mean. If data visualization can change the world, then there is much at stake in the form. The assumption is that access to more data can enable us to make more rational decisions (Dur, 2014). This idea is in part built on the belief in the power of big data to tell us something new about the world (cf. the famous claim that the data themselves are enough and we don't need theories to help us understand them anymore in Anderson, 2008). However, the idea that more data can enable more rational decision-making is deeply problematic. Feminist methodological arguments problematize the idea that research data have intrinsic objectivity (Ramazanoglu & Holland, 2002). Dorothy E. Smith (1974) argues that researchers' claims to objectivity position the researcher as apart from society, as able to take a completely objective viewpoint—what Haraway would call a ‘god trick’ (1988, p. 581). But of course it is not possible to be outside society, and those producing data make decisions which are fundamentally informed by their social positions. This therefore raises important questions about the data that are produced and who is producing them. What assumptions are built in? Who and what is counted? Who and what left out? How do gendered power relations impact the processes of data creation? When we consider ‘big’ data, it is not enough to assume that the data will speak for themselves. We must ask questions of the data (boyd & Crawford, 2012). When it comes to data visualization some research queries the form’s objectivity (e.g. Ambrosio, 2015; Bowie & Reyburn, 2014; Kennedy, Hill, Aiello, & Allen, 2016, and others), but consideration of the political and rhetorical work of data visualizations has been more muted.

In this chapter I draw on my research into online visualizations relating to abortion. On the Persuasive Data project I examine visualizations made by campaigning individuals and groups, and consider how visualizations work in situ as rhetorical devices which attempt to persuade viewers about the rectitude of abortion. My position is pro-choice: I believe that women should have access to safe, legal abortion as a necessary part of healthcare and reproductive rights. For this reason, in analysing these visualizations, it is necessary to think about who is being counted and who is left out, who is doing the data creation and visualization. Drawing on feminist methodological ideas, D’Ignazio and Klein (forthcoming) argue that data visualization has an issue with bodies. They determine that there are four ways in which bodies are missing from data visualization:

1. ‘Bodies are extracted’ (D’Ignazio, Thylstrup, & Veel, 2017, p. 69). States, institutions, and companies have the power to collect data, which means they extract data from people, leaving the ‘bodies’ behind. Institutions determine what kinds of data are collected and what it will be used for, but not all institutions handle sensitive data in a safe, just, or ethical way.
2. ‘Bodies are absent’ (p. 69). The standpoints (including the privileges) of the people doing the work of data creation or extraction and visualization are unacknowledged. This matters because when data are posited to be objective, the privileges and biases of data producers and visualizers are encoded into them without recognition of this fact. D’Ignazio et al. highlight the overrepresentation of white men in tech and STEM: ‘humans might make computers dumber by encoding our age-old biases and structural inequalities into the system’ (p. 69, drawing on Kate Crawford).

3. ‘Bodies go uncounted’ (p. 69). There are differential amounts of data created about things that are important to men and things that are important to women, since, the authors argue, most data scientists are men. For example, there are much more data on erectile dysfunction than on ‘the composition of breastmilk’ (p. 69). The result is that those things on which there are data are seen to be important, whilst those things which are not quantified are not. This produces a very uneven view of the world.

4. ‘Bodies are rendered invisible’ (p. 69). Building on the idea that s/he who makes data and visualizations has an impact on how we see the world, D’Ignazio and Klein argue that visualizations give the appearance of neutrality and objectivity to the data within, whereas, as noted above, they often represent the viewpoints of those who are dominant (see body issues 2 and 3); thus the dominant viewpoints are presented as offering the normal view of the world. Visualizations therefore have discursive power.

Furthermore, O’Riordan (2016) posits that there is a risk of dehumanizing people through the processes of turning us into data and data points, arguing that we need to ‘bring up the bodies’, to re-embody disembodied data. This idea of the missing body is crucial for thinking about what is at stake in data visualization, especially if we want data visualizations to do good work in the world, to change people’s minds, to spur people to action towards making Earth a more just, safe, and beautiful place to be. Fundamentally, data visualization is a creation of people. People—embodied, emotional, enmeshed, messy people—therefore must be at the heart of our critical thinking about data visualization.

My aim in this chapter is to use the four body issues listed above to critically address the work that data visualizations about abortion in the media do in the world. The four issues are not distinct: they interact and overlap with particular results. I also propose that a fifth body issue—that of the viewer—needs to be taken into account if we are to understand what is at
stake in data visualization in the media. Thinking about abortion visualizations in the light of bringing back the bodies enables us to understand why the absence of bodies is a serious problem and how the abstraction of data has the potential to undermine fundamental human rights.

Methodology

First of all, a word on the methodology underpinning the research. In order to understand how visualizations relating to abortion are used by campaigning groups, I used the University of Amsterdam’s Google Image Scraper to scrape Google Images for data visualizations about abortion, whilst also harvesting their URLs for deeper examination. Google Image Search is likely to be a common method for people seeking visual data about abortion. It can therefore be viewed as a valuable tool for groups wanting to influence minds about the rectitude of abortion. In order to gain a sense of what other people may see when using Google Image Search, I cleared my search history to ensure that the results were unaffected by Google’s personalized results system.

The term ‘abortion data visuali*ation’ is most likely to be used by data specialists, but I wanted to get a sense of what kinds of graphical representations of data are available online without being restricted by specialist terms. I therefore also used the everyday alternatives ‘abortion chart’ and ‘abortion graph’. The three terms provided slightly different images, but there was significant overlap, with a number of the same visualizations and the same webpages being returned for each term. I focused on the top 20 search results in each search. These 60 search results are just a snapshot of abortion-related visualizations, but a snapshot has meaning when we acknowledge that few people look beyond a first page of search results. These are the kinds of visualizations that will typically be found and viewed. I paid particular attention to the kinds of data being used, the claims being made in the surrounding texts, and the discourses employed in both written and visual texts. Using these close readings of visualizations in my dataset, I now explore how the body issues can be seen in three of the top visualizations in the results.

Body issue 1: Data are extracted from bodies

One of the major concerns with data visualizations is that, whilst sources of data may be in evidence (i.e. we know who created the data), very little
information about how the dataset was created is usually available. As Bowker (2005) argues, data are never raw, they are always 'cooked'—datasets bear the impression of those who made them. Knowing little about this process is problematic, as organizations may display specific data in particular ways, in order to suit their own agenda. This is the case with the visualizations in my dataset from ClinicQuotes, a US anti-choice blog which gathers together images and stories about the perceived ills of abortion. A large number of the images in the search results come from one page on the blog, ‘Abortion Visual Aids, Graphs and Charts’, which brings together many visualizations and presents them with minimal information about data generation or how data were visualized. One example is the visualization ‘Most Americans say they don’t know enough about the abortion pill to say if it is safe and effective’, shown in Figure 23.1.

The visualization contains two 3D pie charts which show responses to polling about people's opinions about the medical abortion drug mifepristone, undertaken by the Kaiser Family Foundation (KFF). The largest segment of both charts is 'don't know'. The main message of the visualization is that people do not know what to think about mifepristone; they feel ill-informed. Whilst KFF may be supportive of abortion, ClinicQuotes definitely is not, and this visualization used on the site implies support for anti-abortion arguments. The fact of asking this particular question suggests that people ought to be well-informed about mifepristone. But
other questions regarding mifepristone could have been asked. The drug’s safety is arguably not in question since it is approved by the FDA and is regarded as 95 percent effective. In the UK a number of women’s health organizations are calling for medical abortion to be conducted in women’s homes to ensure that they are in a safe environment when they begin to miscarry, rather than travelling from clinic to home. Thus the kinds of questions that could have been asked about medical abortion could relate to the effects of needing to travel to and from clinics and experiencing abortion whilst in transit, for example.

We need to ask questions about the people being polled too: how much is the general population likely to know about the safety and efficacy of any drug? Who was polled? It is likely that the only people qualified to make judgements on the topic are those who are medically trained to evaluate the evidence. Yet the visualization notes only that ‘Americans’ were polled. If the organization were aiming for a representative sample then around half of those polled would be men and a significant number of the women would be post-menopausal, sterilized, infertile, using long-term contraception, or not in heterosexual relationships (Goldstein, 2010). In other words, it is possible that many people polled are unlikely to have much awareness of mifespristone because they are unlikely to come into contact with it. It therefore should not be surprising that more than half the sample said they did not know about the safety and effectiveness of the drug. Ordinary people’s opinions say little about the actual safety or effectiveness of the drug. These polling data should not, therefore, be taken as indicating that it is a problem that people know little about mifespristone, but the visualization shows how particular data questions can be used in order to produce visualizations which reinforce particular political agendas.

**Body issue 2: Visualizers are subject to their own situated knowledge**

Just as the data extraction process is usually opaque in finished visualizations, so is the visualization process. Visualizations are provided as finished products, their clean lines, space, and flat colours drawing on their origins in modernist art (Kennedy et al., 2016). However, like all text producers, visualizers tend to let the beliefs, assumptions, and perspectives characteristic of their own social group—that is, their ‘situated knowledge’ (Haraway, 1988)—influence on the choices they make during the process of production. And data scientists and visualizers tend to be members of privileged groups.
They are often, as D’Ignazio and Klein note (forthcoming), white men. The visualization ‘Abortion Rate & Ratio vs. Poverty Rate’ (Figure 23.2) forms part of a long article about abortion and poverty, and the author, Darwin, presents his views as scientifically based.

The visualization and article encourage the viewer to understand for themselves—to see and know—that there is no correlation between abortion and poverty, and to view these data as the facts of the matter. But there is a relationship between poverty and abortion rates, with poorer women obtaining abortion at higher rates (Jones, Darroch, & Henshaw, 2002) and to deny this obscures the structural reasons for abortion decisions, as I discuss further below, and the continued need for safe access to reproductive healthcare. Darwin is anti-abortion and seeks to bring a scientific examination of data to religious discussions. The article uses the language of statistics, although, note that the timeline on this graph runs backwards, which somewhat undermines the author’s authority when it comes to statistical literacy. The visualization therefore gives a sense of rationality and contributing to informed debate, although there is very little information here. Neither the visualization nor the article discusses why women have abortions, access to contraception, or what it means to be a mother on the breadline, that is, what the actual relationship between poverty and abortion might be.
Both poverty and abortion are taken out of the context of women's lives and decision-making about their families. Darwin suggests that the high rates of recorded pregnancies in 1973 (when Roe v. Wade was passed) represent a euphoric moment: women could now easily get abortions—and so they did. He goes on to argue that numbers of abortions in the US are falling of their own accord due to people realizing that there is a personal cost to terminating a pregnancy. According to Darwin, the fall is therefore a natural decline. Darwin does not take into account that reporting of abortions would have increased post-1973, since abortion was no longer criminalized. No evidence is presented for the claim that the fall in numbers of abortions is due to ‘a build-up of painful experience, which has overcome the initial impression that the costs of getting pregnant (and getting out of getting pregnant) are not as high as they were before 1973’ (Darwin, 2008). Indeed, it is disputed by the UK Royal College of Obstetricians and Gynaecologists (2016), which found that continuing an unwanted pregnancy has a more detrimental impact on women than terminating one. The reasons for a fall in the abortion rate is actually unknown. Thus, Darwin's contribution can be seen as an effort to mobilize data visualization's rhetorical objectivity to support a subjective point of view.

Body issue 3: Data important to women are missing

Very few of the visualizations in my dataset centre pro-choice arguments. Of the 60 visualizations, 28 sit on anti-abortion websites and only 9 are located on pro-choice sites. Others are on news, health, educational (such as university), and visualization critique sites. A large number of anti-abortion visualizations across the dataset (14) are hosted by one site: ClinicQuotes. Anti-abortion campaigning sites use more data visualizations than pro-choice groups, and there is a difference in the kinds of data being visualized. Anti-abortion groups tend to use polling figures relating to opinions on abortion, statistics on numbers of abortions, who has them at which point in their lives and at which point in their pregnancies. On the other hand, the few pro-choice visualizations present charts relating to threats against abortion providers and restrictions on abortions in different states, data on misinformation in state-mandated documents given to women seeking terminations, and visualizations about women's fertility choices over their lifetimes. These offer a different perspective from the anti-abortion statistics. They focus on the tactics of anti-abortion groups and laws in an effort to protect access to abortion. These different topics of pro-choice campaigning
visualizations suggest that the kinds of data relating to abortion that might be useful to women are quite different from the data on numbers of abortions presented by anti-abortion groups. For example, being aware that the information about abortion presented to you by your state has been judged to be misleading (Daniels, Ferguson, Howard, & Roberti, 2016) may enable women to counter the emotive arguments of anti-choice campaigning at the point of decision-making, or it may lessen the emotive impact of such information. However, there remains a gap here in offering data that might be helpful, for example data about how to access abortion in the US (e.g. how far people have to travel to attend a clinic, or how much it costs, or length of waiting times—all things that could be quantified), or the social ‘push’ factors that lead women to conclude that an abortion is the only realistic option. The question of who is socially supported and financially able to raise a child reveals that ‘choosing’ an abortion is not a free choice; it can be a forced decision based on a lack of necessary resources to raise a child, an issue of reproductive justice that has significant intersections with class, race, disability, age, and citizenship status (Lonergan, 2012; Ross, 2017). Data on these aspects of abortion are missing from the examples discussed here.

Since much of the data being visualized by anti-choice groups comes from large statistical polling organizations (e.g. Guttmacher, Gallup), we also encounter the first body issue as well: data are extracted from female bodies for purposes which are not fundamentally about sustaining or extending women’s rights. This becomes more problematic when we think about what the data are that are being visualized, i.e. numbers of abortions, question about the safety of mifepristone. Those data which are visualized come to be seen as important, and those data which are not, to be of no value. That the datasets visualized are created by large well-respected organizations deepens this valuing of particular kinds of data. It raises a further issue of how minimal visualizations strip contextual detail out of issues where such detail is important.

**Body issue 4: Data are abstracted in the visualization**

In her investigation into the use of sonogram images (technical representations of ultrasound data used in examining the foetus inside the womb) by anti-abortion campaigners, Julie Palmer (2009) argues that sonograms have proven highly emotive and powerful tools. In part this is because seeing a sonogram image is confused with knowing the foetus, as if the sonogram provides a real, objective photograph-like image, rather than being
a technological creation. This ‘knowledge’ is then used to further the aim of reducing the time in which women can legally have abortions by making scientific arguments about the viability of foetuses, for example during debate in the UK House of Commons Science and Technology Committee. Those who are experts in interpreting sonogram images acknowledge their ‘beauty’ and emotional power, but contest their ability to tell a truth. They argue that the emotion is in the viewer, not the foetus, and that sonogram images do not produce scientific knowledge in themselves (Palmer, 2009). This conflation of ‘seeing’ with ‘knowing’ is evident in the ‘Abortion Rate & Ratio vs. Poverty Rate’ visualization (Figure 23.2), but presenting data in minimal visualizations as in Figure 23.2 further abstracts both the woman and the foetus, and provides a new layer of perceived objectivity. Using data visualizations could be argued to be a step away from the emotionally arresting images previously used by campaigning groups, e.g. powerfully affective photographs of babies and foetuses (Hopkins, Zeedyk, & Raitt, 2005). However, to see visualizations as only rational, neutral artefacts is to fail to recognize the rhetorical and emotional work that they do. This matters because the abstraction takes abortion out of the context of women’s lives, out of the context of women making decisions that affect them and their families, and that are part of a wider landscape of reproductive decision-making.

This is particularly evident in the Live Citizen visualization, ‘Abortion in the United States’ (see Figure 23.3), which appears on a number of visualization critique sites (the original Live Citizen website has been taken down). What is striking is that bodies are in evidence, but the isotypes and area charts representing data use a widely understood icon for women to tell a political story about women’s place in the world.

The visualization shows statistics about abortion rates worldwide and in the US. It uses metaphors in which the birth rate is represented through visuals of mothering and nursing newborns (women holding babies, prams), and the abortion rate is represented through visuals of women discarding newborn babies into dustbins. Actually, most terminations happen within the first three months of pregnancy when the foetus is not baby-like and could not survive outside the womb. The equation of the foetus with a baby is a common representational tactic in anti-abortion campaigning (Daniels et al., 2016). Blue and pink icons divide the population into equal parts male and female, using the common convention of gendered colour associations. The visualization thus makes use of some common discourses: the gender binary is natural; babies are nursed by women; women are responsible for birth rates and abortion rates; abortion is casually done (the most common reason for
Abortion in the United States

**WORLDWIDE**

- **World Population**: 7 Billion
- **Abortion Occurrence**: 17%, 88%
- **Birth Rate**: 133 Million/yr
- **Abortion Rate**: 42 Million/yr

**UNITED STATES**

- **U.S. Population**: 313 Million
- **Death Rate**: 2.6 Million/yr
- **Birth Rate**: 4.3 Million/yr
- **Abortion Rate**: 1.4 Million/yr

**RACE**

- **U.S. Population**: White 60%, Black 20%, Hispanic 20%, Other 10%

**AGE**

- **15 years old and younger**: 4.5%
- **15-19 years old**: 15.4%
- **20-24 years old**: 23.0%
- **25-29 years old**: 23.4%
- **30-34 years old**: 16.8%
- **35-39 years old**: 8.7%
- **40 years old and up**: 3.2%

**REASONS**

- **Rape or Incest**: 36%
- **Potential Health Problems**: 13%
- **Social Reasons** (i.e., the child is unwanted or unprepared): 8%

**RELIGION**

- **Protestants**: 31%
- **Catholic**: 24%
- **No Religious**: 12%
- **Jewish**: 8.7%
- **Other**: 5.3%

abortion, ‘social reasons’ is described as a child being an inconvenience). This makes for a moralizing tone, reifying women as mothers and demonizing those who terminate a pregnancy. Thus the visualization makes use of data visualization’s perceived objectivity to normalize the responsibility of gestating and raising children as women’s role. This encodes a particular patriarchal viewpoint of gender as biologically given, and of distinct roles for women and men. The data are abstracted and then re-embodied as if they tell the whole truth, but in such a way that distorts.

This brings me to my final missing body problem. Building on D’Ignazio and Klein’s four body issues I determine that we need to think of a fifth group of bodies: those of the viewers of visualizations.

**Body issue 5: The viewer is manipulated**

It is vital that we think about the impact of data visualizations on the bodies of those who view them and beyond: the affected bodies. As my previous research with colleagues on the Seeing Data project (see seeingdata.org) found, visualizations are read in different ways by different viewers, and viewing is influenced by gender, nationality, language ability, education, age (Kennedy et al., 2016), and by the discourses around data, society, and culture (Hill, Kennedy, & Gerrard, 2016). There is no one way to view a visualization, as there is no single way to read a novel: social circumstances change our engagements with culture (Barthes, 1977). However, visualizations do play a role in determining how we read them. The visualization ‘Abortion in the US’ (Figure 23.3) tries to manipulate the viewer to have a strong emotional response against abortion. It does this in part through the ambiguous use of data about abortion, for example through its lack of detail about ‘social reasons’ and baby imagery. As Daniels et al. (2016) have found, providing misleading and inaccurate information about abortion is a key tactic of those who seek to ban abortion, including those who form part of state legislature. The bodies of those seeing visualizations such as ‘Abortion in the US’ may be impacted directly by viewing the visualization; they may find it convincing or upsetting, or have another emotional response (Kennedy & Hill, 2017). Beyond these individuals, however, if data visualizations can change the world, then we need to think about the bodies of those who are impacted at more of a remove. If data visualizations like these are able to change the world, then the direct impact of them may be on women’s ability to access healthcare: working class women who have less money to travel and pay for a procedure; younger women and girls who may be unable to travel; those
who cannot take time from caring, family, or work commitments; black, ethnic minority, and poor women who have less access to contraception; migrant women whose citizenship status means they find it harder to access reproductive healthcare (Lonergan, 2012). Banning abortion does not prevent all abortions, but rather forces women to seek illegal, unregulated, expensive, and often unsafe reproductive healthcare. Combined with the severe restrictions on abortion in some US states, and in the UK a lack of free abortion services for migrant women and those in Northern Ireland (up until 2018), these data suggest that those who are more able to gain abortion services are middle and upper class white women living in the right geographical area and holding the right citizenship. Working class, black and minority ethnic, and migrant groups are disadvantaged through lack of funds and other resources needed to seek out abortion, whether legal, private, or ‘backstreet’ providers. The missing bodies of viewers, and of those who may be affected by decisions informed by visualizations, need to be brought into discussions of the power of data visualizations.

Conclusion

Visualizations about abortion matter when we think about what is at stake in visualizations in the media. These visualizations were in the top portion of Google Image Search results. They have a part to play in changing the world—but not for the better and certainly not because they are providing useful information for making rational decisions. They are offering misleading interpretations of small amounts of data on particular aspects of abortion, leaving out contexts of data creation and visualization, and ignoring significant aspects of factors that affect the experience of abortion. They are leaving the bodies of those from whom data have been extracted, of the visualizers, of women managing their fertility, of women terminating pregnancies, and of those viewing and making decisions based on visualizations out of the frame. These absent bodies mean that it is important to rethink what it means to argue that visualizations can change the world. Abortion is a complex issue and these visualizations and others in my dataset show that simple statistical graphics are unlikely to capture that complexity. But also, more worryingly, such graphics reveal that visualizations are being used as a tool to argue for limits on access to reproductive healthcare. Visualizations can indeed then play a role in changing the world, but it is utopian to imagine that the changes they bring about are always of a positive kind.
References


**About the author**

**Rosemary Lucy Hill** researches gender, popular music, and big data. She is the author of *Gender, Metal and the Media: Women Fans and the Gendered Experience of Music* (Palgrave) and numerous articles on the politics of data visualizations. She is currently investigating sexual violence at live music events.