Data visualisation: Contributions to evidence-based decision-making

By Imogen Robinson, M&E Coordinator at SciDev.Net (@ImmyRobinsonME @SciDevNet)

Click here for the online version

Contents

4.4 Summary of ways in which to enhance the effectiveness of data visualisation

- Appendix A: Methodology
- Appendix B: Research participants
- Appendix C: Focus group questions

Executive summary

SciDev.Net is often regarded to be the world's leading source of news and analysis about science and technology (S&T) for global development. Our mission is to mainstream S&T so policy and practice are better informed and better able to reduce poverty, promote wellbeing and build sustainable economic growth. We do this by publishing news articles, analysis blogs, opinion pieces, features and multimedia pieces that contextualise scientific research and technological innovations within international development.

Recently, SciDev.Net has invested in the production of data visualisation, beginning with a high-profile joint event with the UK Collaborative on Development Sciences on Big data, the open revolution and public engagement. This event investigated the role of data in development and some of the trends, challenges and opportunities around 'data-driven development'. Since then, SciDev.Net has produced a number of data visualisations, including the award-nominated Digital Divide piece, and is building a reputation as a leading source of data journalism for development.

Data visualisation – the visual representation of data in charts and graphs – has grown in popularity in recent years. Media outlets and research communication organisations alike have invested in the production of data visualisation, committing to the belief that visualisation is an effective form of communication.

While a number of claims have been made around the potential of data visualisation as a communication tool, there has been a relative lack of informed discussion around the role that data visualisation can play in the research communication sector.

What is the objective of the report and who should use it?

This report builds on our experiences of producing data visualisations and in data journalism more broadly. It brings together the lessons we have learned through these experiences with insights from the broader sector of research communication.

The primary objective of this report is to encourage a more informed, reflective discussion of the potential and limitations of data visualisation in communicating research evidence. This discussion provides a basis for researchers, research communication managers and journalists to consider the merits of investing in data visualisation production and gain a more informed understanding of how data visualisation can most effectively contribute to research communication goals.

What does the report contain?

Section 1 contextualises the rise of data visualisation and its purported potential to stimulate a 'data revolution' in development. The challenges and opportunities afforded by the increasing availability of data are discussed, and the need for supporting mechanisms which translate data from raw, inaccessible data sets into more meaningful information is explored.

The specific contributions of data visualisation to research communication goals are discussed in **section 2.** Three goals are discussed in this section: selection, engagement and uptake.

1. Selection: the first goal of research communication is to get research outputs to be selected from the pool of competing information available to readers. Data visualisation is well-designed to support selection due to:
    a. Its wide reach
    b. The speed by which it communicates information
    c. The ease by which readers can understand data in a well-designed visualisation

2. Engagement: the second goal of research communication is to get readers to engage with research findings. Data visualisation can support engagement through its interactive capabilities, which enable readers to explore the data. These interactive capabilities, along with the unique visual form of data visualisation, can also prompt engagement by stimulating curiosity in and discussion of the research communicated.

3. Uptake: the ultimate goal of research communication is research uptake: the application of research findings to policy and practice. Again, data visualisation can support the realisation of this goal through its interactive functions, which enable users to explore data in relation to specific questions. Data visualisation can also support uptake through its perceived authority – a key factor in determining research uptake – and by building a culture of data in which the broader public demands greater use of data in decision-making processes.

Section 3 explores instances in which data visualisation is an appropriate form of research communication, recognising that it is not a 'magic bullet' solution to the need for more evidence-informed decision-making, but should instead be used selectively. Instances in which data visualisation can be an effective form of communication are discussed. These instances include:

1. When you have complex, high-quality data
2. When the target audience is able to understand and use visualisations
3. When there is organisational buy-in to the production of data visualisation and a corresponding commitment to the teamwork required to produce such an output

Section 4 discusses ways to enhance the effectiveness of data visualisation. While there are a number of factors that determine the level of effectiveness of a visualisation, research participants identified three as particularly important:

1. The popularity of the subject matter and the timeliness of the visualisation's publication
2. The extent to which the visualisation is tailored to the interests and needs of the audience
3. The quality and integrity of the visualisation's design

Lastly, **section 5**, provides concluding remarks and highlights areas in which further research and discussion are required so data visualisations can be used to the greatest effect in the research communication sector.

Methodology

In addition to existing literature and SciDev.Net's institutional knowledge, this report draws on original research of the broader research communication sector. This research comprises virtual focus group discussions and a series of in-depth semi-structured interviews. Interviewees and online participants were selected through a snowball sampling method[1] and include data visualisation producers, researchers, knowledge intermediaries and journalists. A total of 23 participants from Sub-Saharan Africa, Asia, Latin America, the Middle East, North Africa and the United Kingdom participated in the research, reflecting a range of different viewpoints and experiences. For more detail on the methodology used, please see Appendix A. The report also draws on data collected through SciDev.Net's regular Monitoring and Evaluation systems on the performance of data visualisations.

---

[1] Snowballing sampling is a method by which existing participants identify additional participants from their networks

Limitations of this report

The hypotheses explored in this report are based on the experiences of research participants and supported, where possible, by quantitative data on the performance of visualisations. However, further empirical testing of the hypotheses put forward in this report is required. Where the discussion is speculative, the need for empirical testing is highlighted. This report should therefore be used as a basis for discussion, rather than as a scientific guide to the use of data visualisation.

1. Data and decision-making: Opportunities and challenges

1.1 'Explosion' of data

The 'Digital Age' in which we are purported to live provides us with instant access to a rapidly expanding body of data (Hey 2004; Gregson et al. 2015). Pepper and Garrity (2014, in Gregson et al. 2015: 15) estimate that, in 2014, 90% of the world's stored data was collected in the preceding two years, while Hilbert (2014) reports that the difference in data availability in the late 1980s compared with 2014 was akin to the difference between a layer of newspaper versus four layers of books covering the Earth's surface. This increase in data availability is a result of the proliferation of information gathering and processing capacity which, in turn, is closely linked to the proliferation of technology (Ishmael Perkins 2014).

1.2 Opportunities

The increasing availability of data has led many actors to proclaim its potential to aid decision-making and increase accountability within the development sphere. Data allows us to monitor the outcomes of policies, to monitor national progress towards the Sustainable Development Goals and "to innovate and navigate towards better decisions" (Nigel Shadbolt, in Pearson 2015). Hilbert (2014) explains that analysing data and turning it into information that can help us make predictions could significantly assist development, particularly economic growth. This is because information can reduce uncertainty about the best course of action.

Numerous instances in which data has helped inform decision-making have been identified. Recently, for example epidemiologists in West Africa were able to use mobile phone data which showed people's movements from disease hotspots to predict where new outbreaks of Ebola were likely to occur (Lopes 2015). Similarly, the Tanzanian government is using satellite data to monitor the local spread of malaria in order to revise its policy on malaria containment by targeting the distribution of mosquito nets to areas in greatest need (Ngereza 2015).

1.3 Challenges

While the growing body of available data could enhance accountability and inform decision-making processes, it will not inevitably do so. The mere availability of data is not enough to ensure its use in decision-making. Hilbert et al. (2011) note that, since 1986, the capacity to store information has grown at a much higher rate than our capacity to analyse and communicate that data, indicating missed opportunities. As a SciDev.Net reader remarked in an online survey on data-driven development:

> "*Analysing data in an easy-to-digest way is the biggest constraint [to data-driven development] in my point of view. In a world producing huge amount of data, analysing it to [produce] fruitful information that can support decision-making is a big problem*"

For data to be useful and actionable, investments must be made in analysing that information and communicating it in a way that is "easy to use and practical" (Tatalović 2013). This, in turn, requires investments in mechanisms that enable non-data specialists to understand and use data (Kindornay 2015).
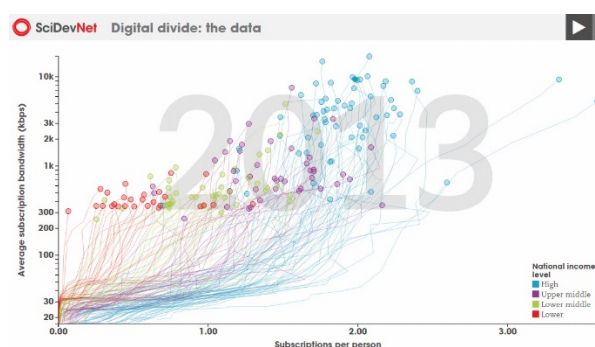
1.4 Data visualisation

One such mechanism for supporting the synthesis, communication and use of data is data visualisation. There is no single agreed definition of the term, but this report defines data visualisation broadly as "the visual representation of data in charts and graphs" (Kennedy 2015). Data visualisations range from simple static graphs and charts to complex interactives, but are united by the function of visualising data (Knezovich 2015).

A relatively 'simple', static data visualisation: A more complex, interactive data visualisation:
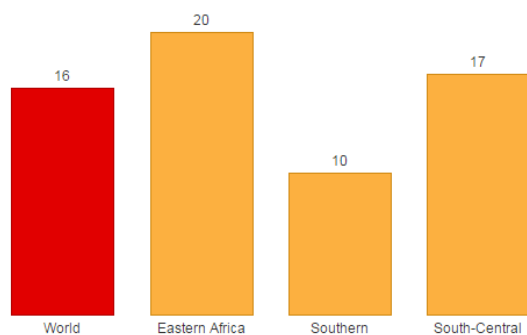
Source: Brown (2016): SciDev.Net   Source: Pollock et al. (2015): SciDev.Net

Data visualisations have gained prominence and popularity in recent years as a result of their perceived potential to render large data sets useful and actionable, and to communicate data in a visually appealing and engaging manner (Global Health Learning 2015; Gregson et al. 2015; Kwapien 2015). Leading media outlets such as the New York Times and the Guardian have made significant investments in growing capacity around data visualisation output in recent years, recognising the value of these products in attracting audiences and maintaining their position as industry leaders (Welsh 2012).

Research institutions and knowledge intermediaries are following the trend, committing to the belief that data visualisation is an effective strategy for enhancing research uptake. SciDev.Net, for example, plans to increase production of data visualisation elements in stories by 60% between 2015 and 2016, while a search of the Overseas Development Institute's site shows a six-fold increase in data visualisation from 2014 to 2015. Additionally, Internews has invested in visualising environmental issues through the InfoAmazonia project and InterAction has established an NGO Aid Map to show the geographical distribution of international development and humanitarian response investments.

While investments and innovations in data visualisation are continually increasing, there has been a relative lack of critical discussion around their capacity to contribute to meeting certain objectives. There is also little thorough exploration of the ways in which data visualisations can help to inform decisions and contribute to forming more evidence-based policies and practices. As the researchers behind the Seeing Data project argue, "in spite of our increasing exposure to visualisations, and our reliance on their ability to tell us 'truths', we don't know much about how people respond to them" (Seeing Data 2016).

Given the growing availability of data, the investments being made in data visualisations and their purported potential to assist in evidence-based decision-making, an informed discussion around how data visualisations can contribute to impact and whether and when they should be used is required. The following sections seek to provide greater clarity around whether and how data visualisations support research uptake so the research communication sector can maximise their potential as a tool to support the application of data in development.
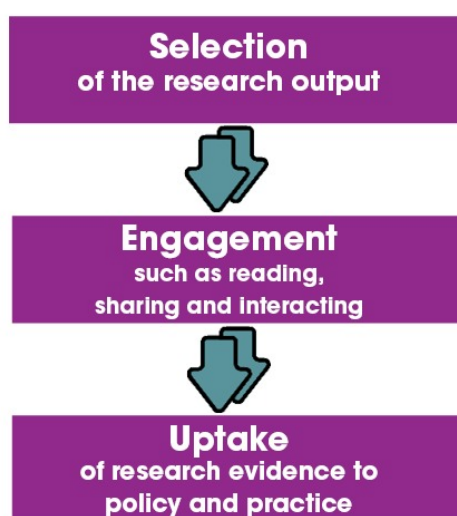
2. What can data visualisation contribute to research communication?

The ultimate goal of knowledge intermediaries such as SciDev.Net is to improve development outcomes by enhancing the application of robust research evidence to policy and practice. This goal is premised on the assumption that policies and practices that are informed by evidence are more effective at reducing poverty, enhancing wellbeing and stimulating sustainable economic growth.

It is widely agreed that policymaking is non-linear (Shaxson 2010), meaning that the application of evidence to policy is not achieved through a single process. Indeed, an initial analysis of the impact cases submitted as part of the 2014 Research Excellence Framework Assessment[2] found that over 3,700 unique pathways from research to impact were reported (King's College London and Digital Science 2015).

While it is clear that research uptake is not a homogenous process, there is a large body of literature around 'sense making', which seeks to shed light on the processes by which users select research, extract information and transform that information into action (see for example Russell et al. n.d.; Abraham et al. n.d.; DFID forthcoming). Within this literature, various stages or steps towards the application of evidence to policy and practice can be identified. These stages include selection, engagement and uptake (cf. Wang 1998; DFID forthcoming). While other stages – such as external checking and validation – can be identified, the following section focuses on these three stages as they are fundamental to the application of research to policy and practice and can be influenced by those who produce or publish research.

Three stages: Taking research and evidence to policy and practice

Selection
of the research output

Engagement
such as reading,
sharing and interacting

Uptake
of research evidence to
policy and practice

---

[2] This is the system for assessing the quality of research in UK higher education intitutions

However, it is important to note that, while data visualisations are well placed to encourage users to select, engage and apply research (discussed below), they are no 'magic bullet' for encouraging or ensuring the application of research to decision-making processes. As Abraham et al. (n.d.) point out, decisions are rarely based on a single research output or piece of information, and instead tend to be based on wider bodies of information, from which users compare and synthesise information, before deciding whether and how to use that information. While data visualisations can help support and encourage greater application of evidence to policy and practice, they are unlikely to offer a panacea for more evidence-based decision-making.

2.1 Selection of research output

The first step towards applying research evidence to policy and practice involves getting 'eyeballs' on research outputs (cf. Wang 1998). With the rise of the internet and the proliferation of technology, online information seekers – including policymakers – face an increasing wealth of information, opinions and resources. Competition to get research outputs noticed is therefore amplified, putting a premium on finding ways to achieve this. Dr Tom Smith, director of Oxford Consultants for Social Inclusion, explains:

> "People are overwhelmed with information at all levels, including decision-makers; we're fighting to get eyeballs on information. Data viz that are well designed and present a clear story are more likely to be taken up than a big report" (SciDev.Net virtual focus group, December 2015)

Data visualisations have various adaptations to help them attract the attention of a wide audience and encourage that audience to 'select' them, including their wide reach, accessibility and speed of understanding. Each of these is discussed in the text below the diagram.

First step to research uptake

The first step is the selection of the research output. Data visualisation contributes to selection because of its:

- Wide reach
- Ease of understanding
- Speed of understanding

1.11 Wide reach

Data visualisations can enhance the likelihood that research will be selected from the pool of competing sources of information because of their wide reach. A number of characteristics result in data visualisations tending to have a wide reach, including:

1. Their visual attractiveness
2. Their 'shareability' on social media

The first way in which data visualisations are well adapted to support selection is through their wide reach. By taking a different visual form to text-based documents, data visualisations are more likely to gain the attention of readers (Amr Elsawy, Egypt). In particular, the aesthetic qualities of data visualisations attract audiences' attention, more so than traditional formats such as text-based

8

outputs, resulting in a wider reach. Prachi Salve, senior policy analyst for IndiaSpend, India's first data journalism initiative, explains:

"Data visualisation helps to break down hard [to understand] numbers into simpler graphs and charts, making it visually appealing for readers. Long, dry and boring stories can be explained in short using interactive charts and infographics"

While there is a lively debate over potential trade-offs between aesthetics and functionality (see for example Lima 2009; McMurray 2009), data on the readership of SciDev.Net articles supports the hypothesis that data visualisations have a wider reach than text-based products. In 2014, SciDev.Net published a range of articles which included data visualisation elements such as graphs and maps. These articles had, on average, 180% more unique page views (UPVs) 30 days after publication than articles published by SciDev.Net without data visualisation elements (see graph below).
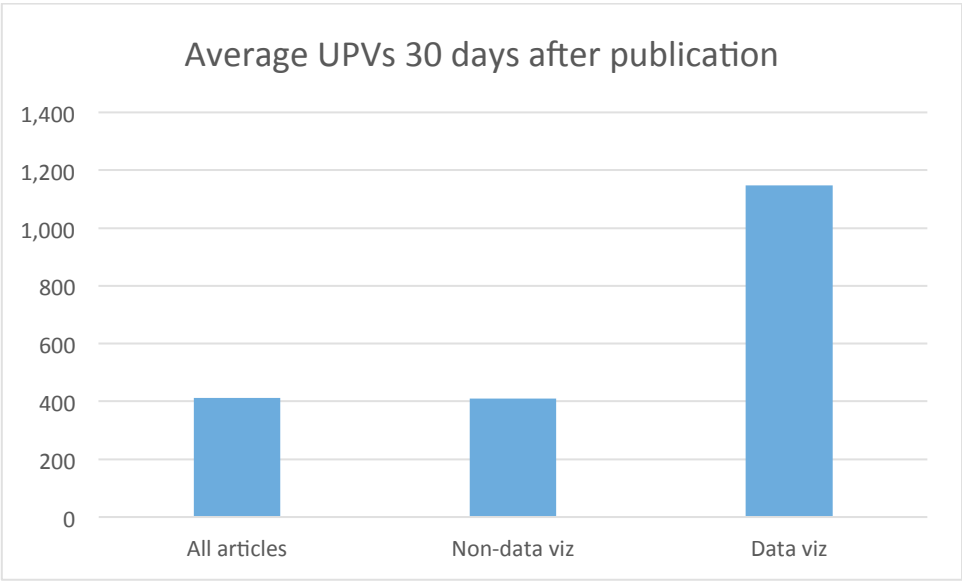
Much of this difference in reach between articles containing data visualisation elements and those that do not is due to social media. Research participants explain that data visualisations are very popular on social media as their visual form makes them attractive to audiences using these platforms. The reach of data visualisation is further enhanced by their 'shareability' on social media platforms.

"Visuals are incredibly important to draw people in from social media" (Anina Mumm, South Africa)

"Data visualisation is hugely shareable, which means it gets disseminated more widely" (Lulu Pinney, UK)

Reaching audiences is a fundamental first step in promoting selection, as audiences cannot select a product they are unaware of.

Data visualisations can also enhance the likelihood that research will be selected from the pool of competing sources of information due to the ease and speed with which they can be understood. Data visualisations can support understanding by changing raw data into patterns and trends

1.12 Speed and ease of understanding

Data visualisation is not only well to reach a large audience, but also to help audiences understand what the visualisation is communicating. In

9

SciDev.Net's 2012 Global Review, over 60% of respondents identified 'lack of sources of information that present S&T information readily usable for public engagement' as a major challenge in engaging readers on issues related to S&T (Romo 2012: 35). Data visualisations help to address this gap by presenting data in a way that is easier to understand.

Kwapien (2015) explains that data, in isolation, is meaningless. To the untrained eye, data housed in large spreadsheets is inaccessible, unengaging and unappealing. As Surendran Balachandran of SocialCops, India, notes:

> "Excel sheets … aren't everyone's friend"

Data visualisations enhance accessibility of data by changing the form of data from raw, unprocessed, discrete packets into key patterns and relationships ('information') (Ackoff 1999; Hey 2004; McCandless 2010). In doing so, audiences' capacity to understand the data is enhanced.
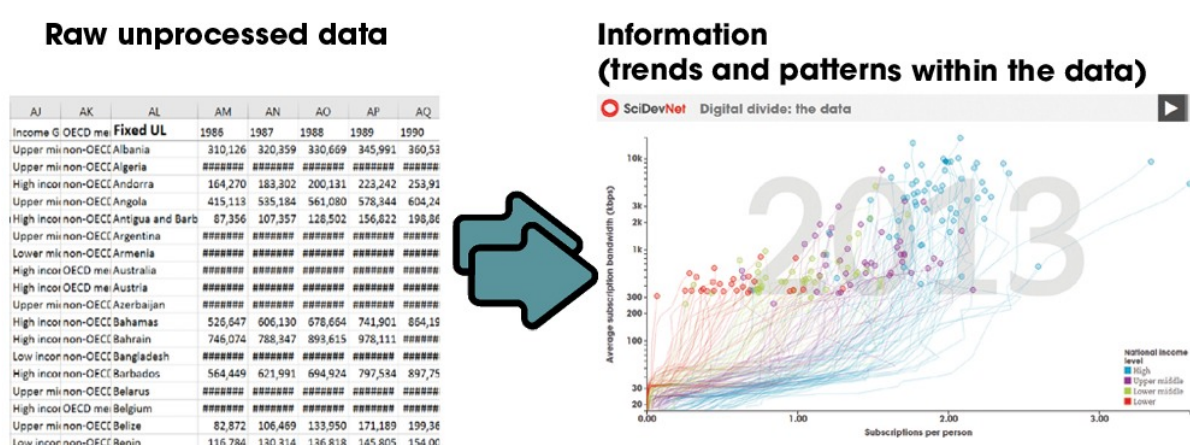
Source: Martin Hilbert and Pollock et al. (2015)

Dr David Tarrant of the Open Data Institute explains that the brain finds it easier to recognise and process patterns and trends than numbers, meaning products that visualise trends and patterns within data, as opposed to whole data sets, are more meaningful and accessible for audiences and can be understood more quickly than text and numbers (see also Few 2014). This is because data visualisation

> "appeals to the dorsal stream in our occipital lobe … The dorsal stream is one of the fastest thinking parts of the brain … One picture, shown for fractions of a second, is enough to trigger a lot of reaction in the brain … Text, audio and video require a lot deeper, and slower engagement, to potentially tell the same message"

The summative function of data visualisation was identified as a key comparative strength of data visualisations by a number of respondents:

> "We're competing for people's attention more than ever before, so visual messages may work better than reading lots of text, particularly if it's filled with complex numbers and data" (Anina Mumm, South Africa)

> "Readers could … use the visuals to get a snapshot of the article's content without necessarily reading the entire piece" (Gibert Nakweya, Kenya)

Identifying methods by which information can be articulated quickly is of particular importance when targeting policy audiences. As is well documented (DFID forthcoming), policymakers are time-poor and impatient in their pursuit of evidence. Visualising information that highlights key trends and relationships within data in a form that the brain can process quickly enhances the likelihood that information will be successfully communicated. Research participants explain:

> *"Data viz is particularly useful for communicating information quickly to people with limited time"* (David Girling, UK)

> *"In [South Africa], the easier it is to get the message across, the quicker it will be to obtain buy-in from all stakeholders, including policymakers and funders"* (Shirona Patel, South Africa)

> *"People have a short attention span,[so] the quicker the message can be conveyed the better"* (Lebo Majara, South Africa)

By reaching a large audience and supporting that audience in quickly understanding the data being communicated, data visualisation is well designed to prompt audiences to select the visualisation from the pool of information competing for readers' attention. Done well, data visualisation can therefore be an effective format by which to get your research seen, selected and understood by a large audience.

2.2. Engagement with research output

The next step towards research uptake, after selection, is to get readers to engage with the research output. Engagement is an important step as it encourages readers to assimilate the information being communicated and to situate that information within their existing frameworks of knowledge. Gilbert Nakweya, a freelance journalist from Kenya, explains:

> *"Data visualisation helps us to construct meanings from research work and get to know what it practically implies [for] us. We often have frameworks through which we construct and view our world, data visualization helps us to place pieces of information within our frames to construct meaning"*

Engagement can take a variety of forms, including sharing, saving, commenting and 'liking' (cf. DFID forthcoming). Visualisations are well-placed to encourage readers to engage with research, as demonstrated by a range of webmetrics. In section 2.11, it was shown that articles published by SciDev.Net which contain a data visualisation had 180% more unique page views 30 days after publication. Similar trends can be observed in other engagement metrics. For example, articles containing data visualisation elements had 533% more retweets than articles without data visualisation elements (see graph below).

Average retweets 30 days after publication
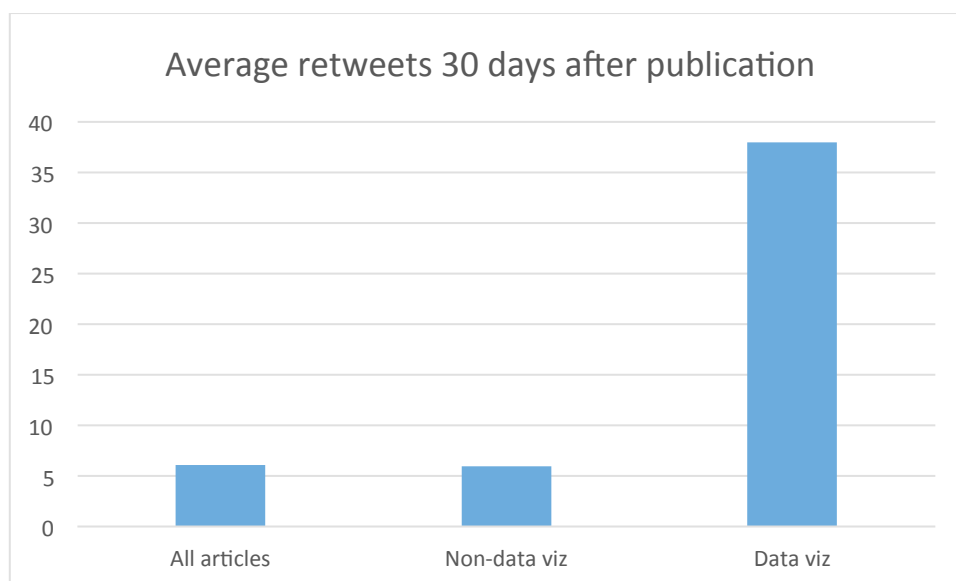
There are a number of ways in which data visualisations are well suited to encourage engagement, including by facilitating interaction and prompting discussion.

Second step to research uptake

The second step is audiences engaging with the research output. Data visualisation contribute to engagement through:

- Interactivity
- Stimulating curiosity
- Fostering discussion

2.21 Facilitate interaction

Data visualisations can support engagement with data by enabling the reader to interact with that data.

The first way in which visualisations can support and encourage engagement with data is through the provision of interactive functions. As a result of software developments, data visualisations are becoming increasingly innovative. In particular, data visualisations with interactive capabilities are becoming more common, enabling users to explore and experiment with the data portrayed. Such innovations meet an expectation for greater interactivity, which is partly driven by the proliferation of touch-screen technology (Knezovich 2015), and encourage readers to engage with data presented in a way that is impossible in more traditional text-based outputs.

Caleb Gichuhi, communications officer for The Institute of Social Accountability in Kenya, explains:

> *"Animated infographics are more successful in stimulating curiosity; getting users to explore the issue more in depth"*

Data on the extent to which audiences make use of interactive capabilities is difficult to obtain, but analysis of SciDev.Net's The hidden digital divide interactive[3] published as a stand-alone interactive by Dawn Pakistan shows that 91% of unique sessions had an 'event' of some kind (e.g. the 'play' button was clicked or the scatter graph function was used), indicating that readers make use of interactive capabilities.

2.22 Stimulate curiosity

> Data visualisations can also prompt engagement by stimulating curiosity in data.

Secondly, data visualisation can encourage users to engage with data by stimulating interest in that research. By attracting attention through their unique visual form and interactive capabilities, visualisations can prompt curiosity in data on issues that do not typically gain much attention from the public or target audience. Caleb Gichuhi of the Institute of Social Accountability in Kenya explains:

> *"The fact that a [visualisation] can act as a summary of complex info and can … generate curiosity on otherwise 'boring' info is a major strength. I have had people ask me for a full 40-page document only after going through a one page [visualisation] citing that the [visualisation] made them curious about the full document from which it was created and that it made them understand a situation which they had no clue about"*

The visualisation to which Gichuhi refers provides information about the budget and county assemblies in Kenya, which *"a lot of youth here find boring",* but in which interest was stimulated through the production of an engaging visualisation.
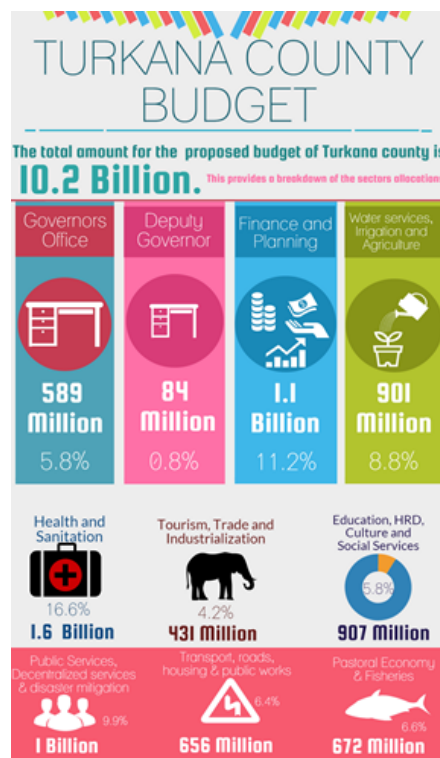
Source: The Institute of Social Accountability

---

[3] This data relates to the piece as republished as a stand-alone visualisation by Dawn as opposed to the longer form article published on SciDev.Net, for which data on events is not available.

2.23 Prompt discussion

Data visualisations can also support engagement with data by prompting discussion. A number of characteristics of data visualisations make them well-situated to prompt discussion including:

1. Interactive capabilities
2. Popularity on social media

As well as stimulating interaction and curiosity, data visualisation is well suited to prompt discussion. Lulu Pinney, author of Telling Information, explains:

"People can interpret [data visualisation] through their own filter. You can interpret based on your own starting point. This [scope for interpretation] opens the door to discussion"

This propensity towards discussion is amplified by the success of data visualisation on social media, where discussions tend to be livelier. Mott MacDonald, a consultancy firm that led an evaluation of online portals funded by the UK Department for International Development (DFID), explains: "With social media, the internet makes information sharing 'real' and real time: personal views and information can be shared and immediate feedback reinforces the sense of debate and dialogue" (DFID forthcoming). Through its popularity on social media and its propensity to drive discussion, data visualisation encourages readers to engage with the data presented.

Again, data collected at SciDev.Net supports this idea that data visualisation encourages readers to engage through discussion: articles containing a data visualisation element had an average of 942% more comments on Twitter compared with those without such elements.
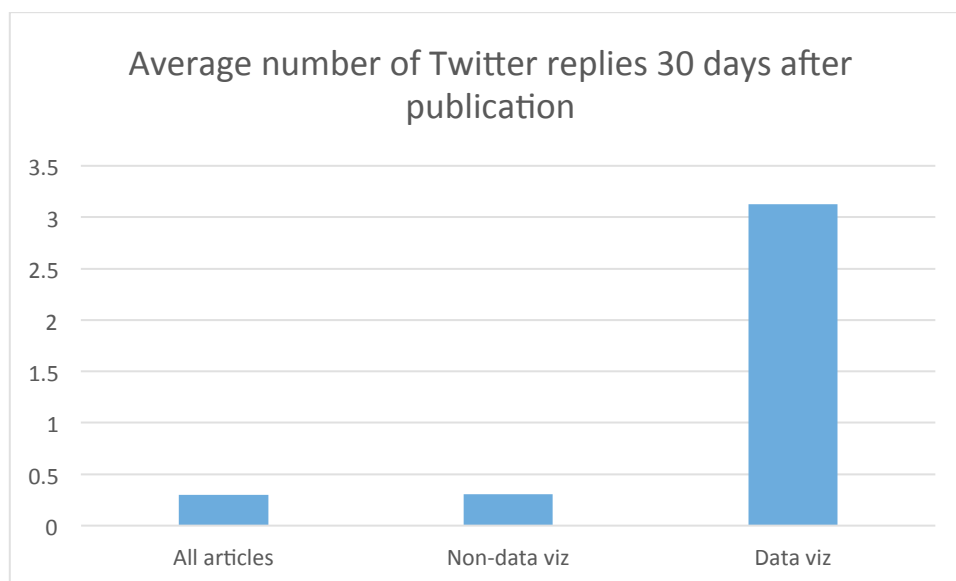
Drilling down into this data, it becomes apparent that interactive data visualisations stimulate more comments than non-interactive data visualisations. In the sample discussed above, articles containing an interactive data visualisation element received 1,341% more comments than those that contained a non-interactive element, suggesting that interactive data visualisations are particularly well suited to stimulating discussion due to the scope for interpretation that they offer.

Average number of Twitter replies 30 days after publication

A bar chart titled "Average number of Twitter replies 30 days after publication" with a y-axis from 0 to 6. "All articles" is approximately 0.3, "Non-interactive data viz" is approximately 0.35, and "Interactive data viz" is approximately 4.8.
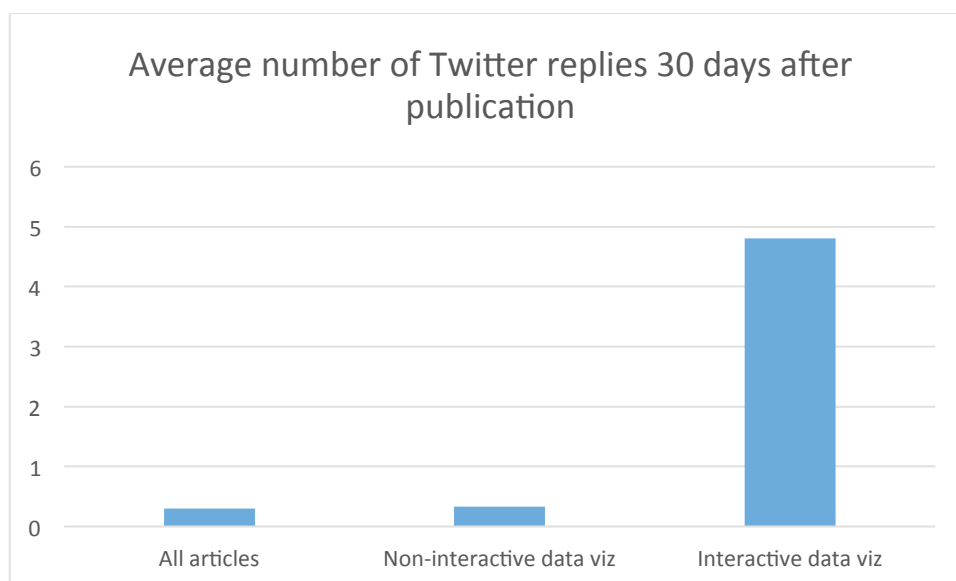
Caveat: the importance of good design

While data visualisation can enhance engagement with research outputs, it is important to note that a poorly designed data visualisation can result in lower levels of use than a text-based counterpart. For example, visualisations that have poorly labelled axes, that feel 'overwhelming' or that are difficult to understand quickly can prompt the user to 'skip over' the visualisation, meaning the output is neither selected nor engaged with by users. Dr Zipporah Ali, director of the Kenya Hospices and Palliative Care Association, warns:

> *"Visualisations need to be easy to understand, not too complicated. Sometimes data visualisations are not easy to understand and your eyes get tired [so] you move on"*

All research participants emphasised the importance of following good design principles when producing data visualisations (discussed in more detail in section 4). This is vital to unlocking the potential of data visualisation to prompt readers to engage with research.

2.3 Uptake/application of research evidence to policy and practice

While selection and engagement are valuable intermediate outcomes of data visualisation, the ultimate goal of research communication is to promote research uptake. DFID defines research uptake as "research findings being used in international decision-making, such as by policymakers or practitioners" (in DFID forthcoming). Prompting the application of research evidence to policy and practice is a notoriously difficult goal to achieve (Shaxson 2005), but data visualisation can support the step from selection and engagement to research uptake in a number of ways, including through interactivity, authoritativeness and by helping to cultivate a 'culture of data'.

Third step to research uptake

The third step is the application of research to policy and practice. Data visualisation contribute to the uptake/application of data through:

- Interactivity
- Authority
- Building a culture of data

2.31 Interactivity

Data visualisations can support research uptake by enabling users to interact with the data. This helps readers to:

1. Gain a deeper understanding of the meaning of the data
2. Answer questions that are specific to their work or context

In section 2.21, the propensity of interactive data visualisation towards greater engagement was discussed. As well as encouraging readers to engage with the data presented, interactive visualisations can help readers to practically apply that data to real-life decisions.

Data is commonly perceived to be abstract; disconnected from the 'real' world. This disconnect means the implications of data to real-life decisions can be unclear. Dr Zipporah Ali, director of the Kenya Hospices and Palliative Care Association, explains:

> *"Figures and charts can be very impersonal … It is not always clear what the data means or why it matters to real life"*

Indeed, a lack of supporting mechanisms that take data from the confines of the research setting and into the real world is an oft-cited barrier to research uptake. In SciDev.Net's 2012 Global Review, for example, over 50% of respondents cited lack of social, political or economic analysis as a challenge to applying information about science and technology to development-related activities (Romo 2012).

Data visualisations can help readers feel 'closer' to the data and gain a better understanding of the meaning, implications and applications of that data (Shneiderman in Singer 2011). Anina Mumm, science communication and digital media specialist at ScienceLink and SciBraai South Africa, explains:

> *"Interactive data viz … opens up a whole new world of personalized storytelling, such as location-specific data that allows users to zoom in on stories in their location on a map"*

This notion of 'personalised storytelling' is elaborated by Hassel Fallas, data-driven editor at La Nacion, Costa Rica:

> *"There are stories that, given the volume of data involved, are impossible to be told in a text; but if we have an interactive display we can generate information and knowledge at a glance. [This] makes available to people the conclusions based on data analysis that do not*

only show the conclusions of the journalist, but also allow people to interact, find and create their own story"

Interactive elements enable users (including policymakers) to explore data in relation to specific questions. This is important as the needs of policymakers are "pragmatic, rather than academic": policymakers need to know what policies have worked in other situations, and whether their success can be replicated in their own constituency (DFID forthcoming).

By enabling readers to interact with data, visualisation helps to increase their meaningfulness and utility. Lima (2009) explains that, by enabling users to select specific data, play around with different variables or focus on certain geographical areas, interactive data visualisations can help readers identify and explore questions or issues that are pertinent to them.

At SciDev.Net, we have used interactive capabilities in a number of our visualisations. For example, we produced an interactive visualisation of a data set comparing the performance (in terms of total yield in kilograms per hectare) of different bean types in two districts of Uganda. The data set and the resulting visualisation offers a large amount of information which, to the reader, may feel overwhelming.

Bean variety yield in two districts of Uganda, 2013
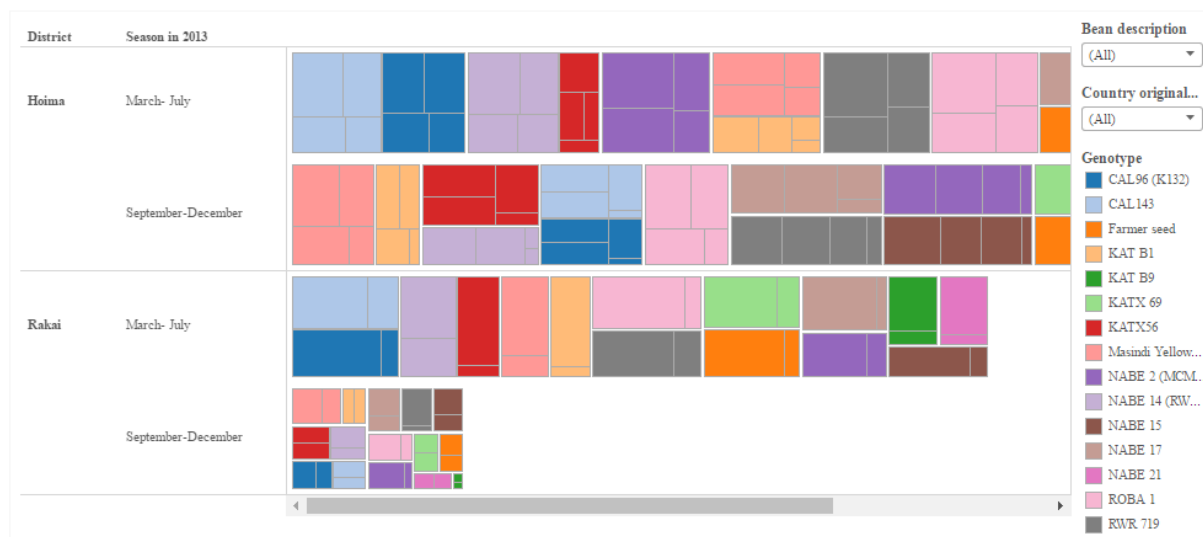
Source: Awio and Vesper (2016)

Adding interactive capabilities, however, helps the reader make sense of the data. For example, a 'hover' function enables readers to easily access more information about each of the data points – in this case to get information about the origins of the bean and its yield.
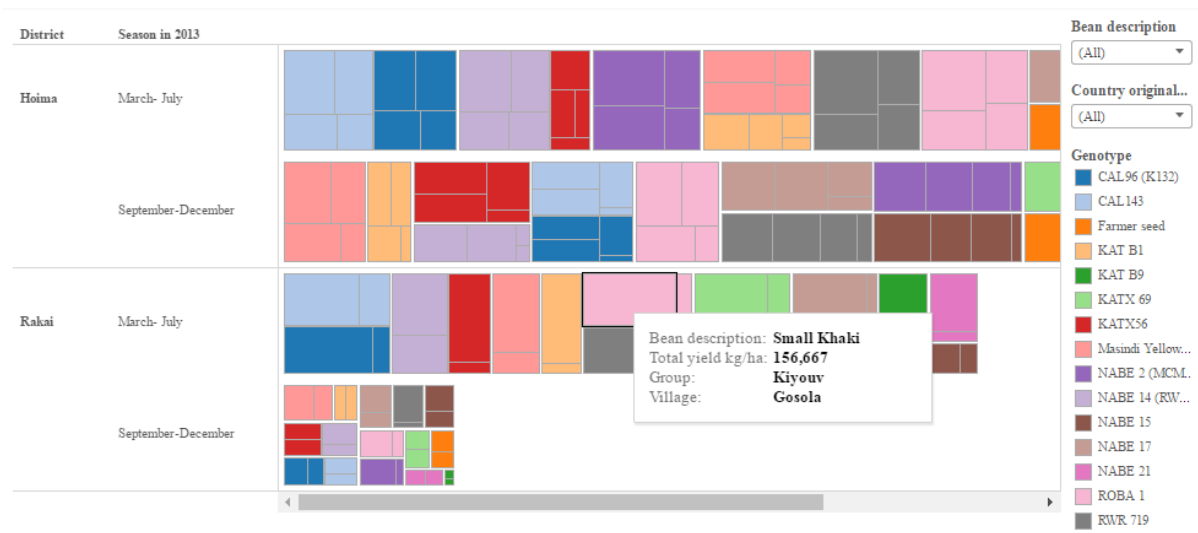
Adding filter options enables users to narrow down the visualisation so that the information portrayed is relevant to them. For example, a farmer in the Hoima district trying to decide what bean to farm in September-December could click on their district and the relevant season to narrow down the data shown in the visual.
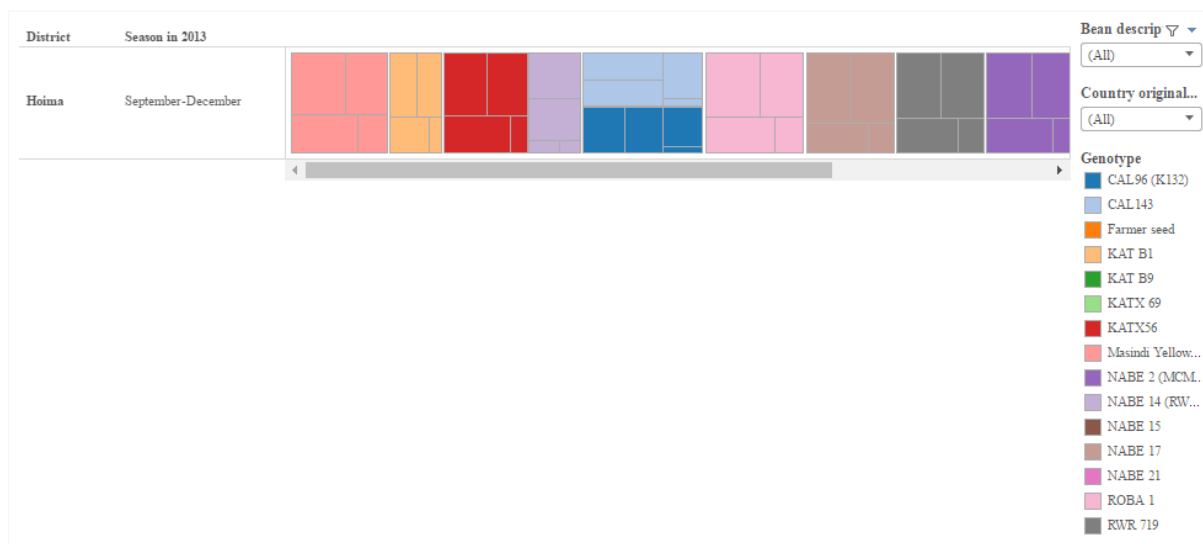
The farmer could then use the specialised filter function (on the right-hand side) to filter the data further to compare yields of a specific type of bean – for example the large red kidney bean which may be preferred by customers in that district – and select the bean that, historically, has had the highest yield in that season and that district.
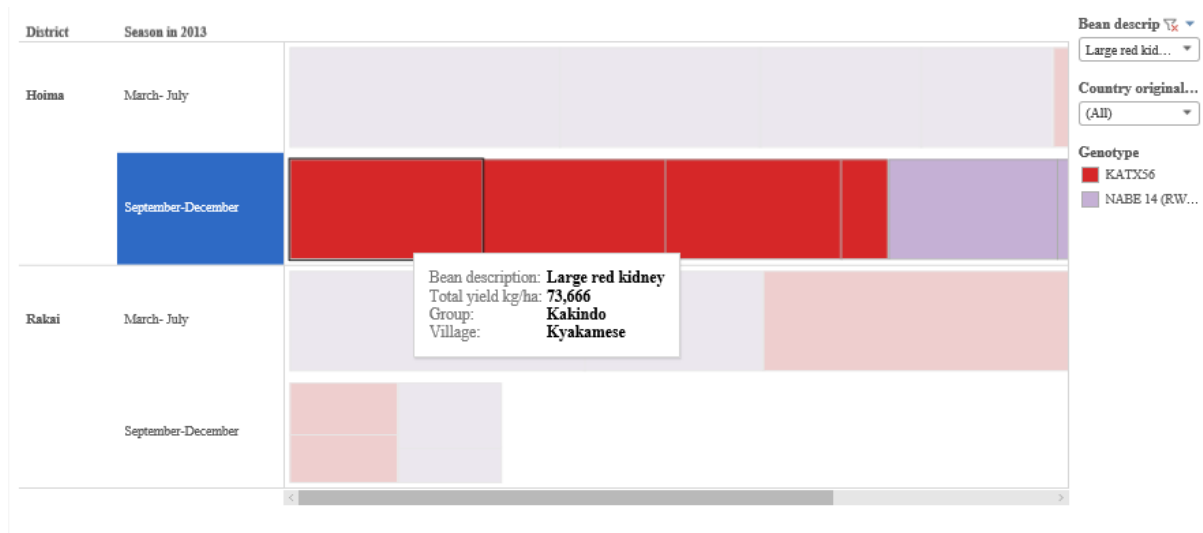
Awio and Vesper (2016): SciDev.Net

Adding interactive capabilities to data visualisations that contain a large amount of data and a large number of variables enables the user to reshape the layout of the visualisation and 'drill down' into the data to answer questions that are specific to their situation (cf. Lima 2009). This then enables users to

> "construct meanings from research work to get to know what it practically implies [for] us" (Gilbert Nakweya, Kenya)

By bridging the gap between 'abstract' research and real-life decisions, interactive data visualisations are well adapted to promote and support research uptake. As Haque 2014 explains: "One of the best ways to make data more meaningful is to make it yourself." While it is impractical for every policymaker and practitioner to go out to collect data and conduct empirical experiments, enabling readers to experiment by changing parameters from already collected data helps to enhance the

---

Caveat: the need to be selective in adding interactive functions

However, while adding an interactive function can enhance audience engagement with data, it is important to remember that having an interactive component is a decision, rather than a necessity. Andrew Lee, head of digital at SciDev.Net, explains:

> "Interactivity is not always needed. If you just want to represent differences on a map, it does not need to be interactive … You need to think about whether interactivity will add value to the end user"

As the next two sections indicate, interactivity is not the only way in which visualisations can promote uptake.

---

meaningfulness and utility of that data and the likelihood it will be acted on.

2.32 Authoritativeness

> Data visualisations can also support research uptake through their perceived authoritativeness – a factor found to be key in determining research uptake.

Data visualisations also support research uptake through their perceived authoritativeness. Over 85% of respondents in our Global Review identified 'authoritativeness' as a valuable attribute that helps support research uptake (Romo 2012). Authoritativeness of research is closely linked to 'trustworthiness' and involves credibility, reliability, reputability and accuracy (Romo 2012). The importance of the perceived authoritativeness of research was also identified as a key factor in research uptake by a systematic review of barriers to and use of evidence by policymakers (Oliver et al. 2013).

Data visualisations are commonly viewed as authoritative sources of information due to their association with 'factual' information. Kwapien (2015) explains: "Stakeholders from different industries have already realized that data visualisations are more likely to attract attention and gain trust than other facts as they carry an aura of seriousness, even when they are designed to mislead" (see also McMurray 2009; Cohen n.d.). Gilbert Nakweya, a science journalist in Kenya, has noticed a similar trend, remarking:

> *"Data visualisation helps newspaper outlets and the media to capture public trust and confidence … compared to text-based articles, data visuals make news stories more credible and widely read"*

This sentiment is supported by data collected by the Aberdeen Group, which found that 64% of interactive visualisation users and 39% of static data visualisation users had improved trust in the data on which the visualisation was based after viewing/using the visualisation (Krensky 2014). This data was derived from a survey of 676 data visualisation users – 227 of whom used interactive

---

Caveat: the need to be selective in adding interactive functions

However, while adding an interactive function can enhance audience engagement with data, it is important to remember that having an interactive component is a decision, rather than a necessity. Andrew Lee, head of digital at SciDev.Net, explains:

> *"Interactivity is not always needed. You need to think about the story you are trying to tell and what features will add value for the end user"*

As the next two sections indicate, interactivity is not the only way in which visualisations can promote uptake.

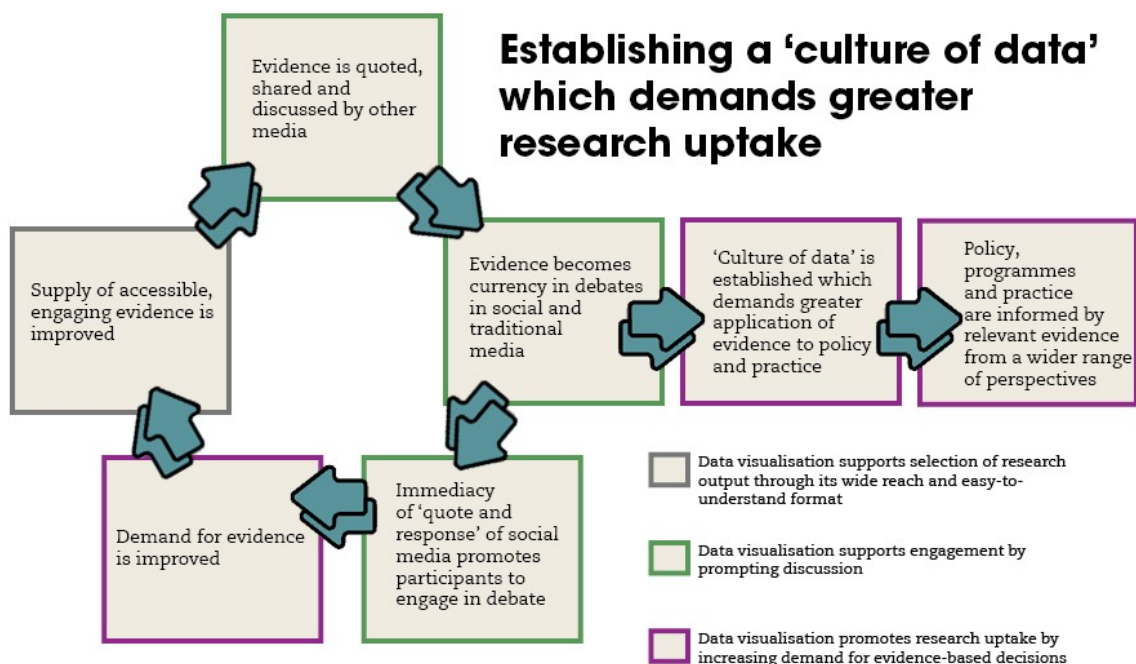visualisations and 449 static visualisations.

2.33 Culture of data

Thirdly, data visualisations can support research uptake by helping to establish a 'culture of data'. They can do this by making the public more confident in understanding and using data which, in turn, can stimulate greater demand for the use of data in decision-making processes.

Finally, data visualisation can prompt greater application of data to development indirectly by contributing to the creation of a 'culture of data'. Srikanth Viswanathan, of the Janaagraha Centre for Citizenship and Democracy in Bangalore, explains:

"A secondary purpose [of data visualisations] is to create a culture of data and make a wide variety of stakeholders comfortable with data. Data visualisation helps accomplish this by softening the impact of numbers and analysis which some stakeholders may be unfamiliar with or intimidated by"

By increasing the accessibility of data (section 2.1), visualisations help readers become more 'comfortable' and engaged with data (section 2.2). Data therefore becomes 'currency' in public discussions and debates, creating a sort of culture of data in which the utility of data to decision-making processes is better appreciated by the public, leading to greater demands for data to be applied to decisions around policy and practice (cf. DFID forthcoming). Thus, the contributions of data visualisation discussed in this section – selection, engagement and uptake – come together to stimulate public demand for more accountable, evidence-based decision-making.

Establishing a 'culture of data' which demands greater research uptake

Evidence is quoted, shared and discussed by other media

Supply of accessible, engaging evidence is improved

Evidence becomes currency in debates in social and traditional media

'Culture of data' is established which demands greater application of evidence to policy and practice

Policy, programmes and practice are informed by relevant evidence from a wider range of perspectives

Demand for evidence is improved

Immediacy of 'quote and response' of social media promotes participants to engage in debate

Data visualisation supports selection of research output through its wide reach and easy-to-understand format

Data visualisation supports engagement by prompting discussion

Data visualisation promotes research uptake by increasing demand for evidence-based decisions

Adapted from DFID forthcoming

By helping to build a body of data users who believe in the importance of data and demand its use in decision-making processes, data visualisations promote the greater application of data to policy and

---

Caveat: the lack of empirical evidence on research uptake

As noted earlier in this report, policymaking is a non-linear process that is influenced by a number of different factors (Shaxson 2010). Because of the complex nature of the policymaking process, hypotheses about research uptake are difficult to test. This section has suggested ways in which data visualisation can promote research uptake, based on research participants' observations and existing literature, but it should be noted that these are largely hypothetical, rather than empirically proven.

---

practice. Indeed, Parks (2014) identifies such public demand as a requirement to establish a 'data revolution' in which accountability and transparency in decision-making processes is enhanced through the application of data.

2.4 Summary of the contributions of data visualisation to research communication

First step to research uptake	Second step to research uptake	Third step to research uptake
The first step is the selection of the research output. Data visualisation contributes to selection because of its: • Wide reach • Ease of understanding • Speed of understanding	The second step is audiences engaging with the research output. Data visualisation contribute to engagement through: • Interactivity • Stimulating curiosity • Fostering discussion	The third step is the application of research to policy and practice. Data visualisation contribute to the uptake/application of data through: • Interactivity • Authority • Building a culture of data

This section has discussed the contributions of data visualisation to three major research communication goals: selection, engagement and uptake. The wide reach, accessibility and speed of understanding of data visualisation were described as key attributes in encouraging readers to select visualisations from the pool of competing information with which readers are faced and to get more 'eyeballs' on the research findings. The interactive capabilities of data visualisation were discussed in terms of stimulating curiosity in and prompting discussion around research findings and, in doing so, enhancing readers' engagement with research. Finally, the interactive capabilities of visualisations were also described as facilitating research uptake by helping to address the pragmatic needs of policymakers and practitioners. The perceived authoritativeness of data visualisation was also found to be a key factor in promoting research uptake, while the capability of data visualisation to generate a 'culture of data' in which the public demands greater application of evidence to policy was also discussed.

3. When is data visualisation an appropriate form of research communication?

While in many instances data visualisations have significant potential to enhance the accessibility, utility and applications of data, visualisation is not always the most appropriate way in which to tell the story of a data set. Dr David Tarrant, senior trainer at the Open Data Institute, warns:

> *"Just because we can visualise data (and yes you can visualise the majority of it) doesn't mean we should"*

A similar warning is articulated by Dr Tom Smith, director of Oxford Consultants for Social Inclusion:

"Data viz are not always the most appropriate communication tool. Sometimes a good analogy may be more effective"

Data visualisation is one tool in the research communication toolkit, and investment in the production of data visualisation should be selective and carefully thought through. It should not be assumed that data visualisation is the most appropriate form of communication. For example, if the purpose is to prompt an emotional reaction from the audience, other forms of communication can be more effective. Dr Zipporah Ali, director of the Kenya Hospices and Palliative Care Association, explains that telling the story of an individual's suffering using tools that lend themselves to emotive storytelling, such as video, can be more effective at gaining support for the association than presenting 'impersonal' data. Shirona Patel, head of the communications department in the School of International Development at the University of Witwatersrand in South Africa, makes a similar comment:

"We are … working on a project where we need to explain the comparative benefits of a sugar tax on sugar sweetened beverages to policymakers and the public. We need to convey the tax implications and the benefits (in life years) of the proposal using data visualisations, which is proving to be much more difficult [than we thought]. It is probably easier to convey the benefits using a 'softer more emotional' form of communication – video, audio rather than data. We have not found the right mix yet"

To be effective, the method of communication must be tailored to the message being told and the audience being targeted. The appropriateness of data visualisation will therefore vary depending on the subject matter as well as audience characteristics, meaning that data visualisation will not be an appropriate form of communication in all instances. Fatou Gueye, academic support officer for the Education and Research in Agriculture project in Senegal, for example, warns that:

"In Senegal … the best [communication] strategies are on TV and mobile phones [and] poster campaigns"

With this in mind, the following section discusses instances in which more complex, interactive visualisation is an appropriate form of research communication and a logical investment to make. This section focuses on so-called 'next generation' visualisations (i.e. interactive visualisations) as opposed to static visualisations, as the former generally require more significant investments – both in terms of time and money – than the latter, which can typically be quickly produced using readily accessible programmes such as Excel (cf. Gatto 2015).

The following section discusses three instances in which data visualisation can be appropriate for research communication: 1) when you have high-quality data which you want the user to explore, 2) when the audience has the capacity to effectively understand visualisations, and 3) when your organisation has the capacity to produce visualisations.

3.1 High-quality, complex data

For data visualisation to be an effective form of research communication, you must have a high-quality data set. Data visualisation can be particularly appropriate where the data set is complex.

Decisions about investments in data visualisation must begin with the quality and availability of data. Unless accurate and complete data is available, the visualisation will lack integrity and authoritativeness and could misinform readers. Hassel Fallas (Costa Rica) explains:

"If we are working with data, it should be done

with rigour and accuracy, verifying the data [just] as [with] any other source. We need to ensure the precision and trustworthiness of data that will be presented as a fact to our readership"

Availability of high-quality, accurate data is a major barrier to the production of data visualisation, in spite of broader trends in growing data availability (section 1). Research participants explain:

"As a designer, I'd say the main challenge is the quality of the data. Ensuring data is validated and suitable for communicating can be a challenge. It is important that everyone in the project realises the importance of having accurate data" (Ricci Coughlan, UK)

"Decision-makers rely on accurate data to make big decisions and if the data is flawed in any way, then no matter how well executed the [visualisation] is, it will not be taken seriously" (Calen Gichuhi, Kenya)

Even if high-quality, accurate data is available, data visualisation is not necessarily the optimal form of research communication. For example, where a data set tells an uninteresting or unsurprising story that could be communicated in a few sentences, investment in a data visualisation is unlikely to pay off (Dr Tom Smith, UK).

However, where an interesting and high-quality data set exists, visualisation can be an appropriate form of research communication. In particular, data visualisation can be an optimal format when the data set is complex (e.g. contains a number of different variables showing a number of patterns and trends). In this instance, an interactive data visualisation which enables the user to sort, filter and explore the data can be appropriate as it enables more information to be communicated in limited space compared with text (Gatto 2015), as demonstrated through the discussion of SciDev.Net's visualisation of bean yields in two districts in Uganda (section 2.31).

3.2 Audience capacity

Audience capacity is also crucial in determining whether data visualisation is an effective form of research communication, because visualisations require users to have certain skills.

The appropriateness of data visualisation as a communication tool also depends on audience capacity, for data visualisations demand certain skills of the reader. 'Graphicacy' (the ability to understand and use a map or graph) is often taken for granted, but not all audiences have the same degree of familiarity with or understanding of graphs and charts, particularly where such graphs and charts use advanced techniques (Global Health Learning Center 2015; Kennedy 2015). Research participants explain:

"Utility of data visuals varies from [one] audience to another. For instance, education levels of audiences determine to what extent one can utilise a combined line and bar graph, for example" (Gilbert Nakweya, Kenya)

"Audience and mass communicators in some countries are not yet used [to] infographics, their eyes are not used to the keys of the colours and the keys of the maps, and the different kinds of charts. So the challenges for the graphic designer [are] to simplify the design and use [a] 'step by step' way to teach the audience how to deal with infographics" (Amr Elsawy, Egypt)

The need to align data visualisations to audience capacity has been highlighted in the Seeing Data research project, led by Professor Helen Kennedy. Through focus group research, interviews and diary-keeping which assessed how readers engage with data visualisations, Kennedy (2015) found

that confidence and skills related to things such as language, statistics, visual literacy, computers and critical thinking all affected users' engagement with data visualisations. Where visualisations were not aligned to the audiences' capacity, engagement with data visualisations used in the research was low. This was demonstrated through participants' reactions during Kennedy's research:

> *"It was all these circles and colours and I thought that looks like a bit of hard work; I don't know if I understand"* (Sara, 45, a part-time careers advisor)

Decisions about the appropriateness of data visualisation as a form of communication must therefore be guided by the capacity of the target audience to understand and use the functions of the visualisation. As with any form of communication, it is imperative

> *"to keep the end user in mind at all points"* (Andrew Lee, UK)

Kennedy (2015) also found that the time the audience has available to explore the visualisation affects levels of engagement and has implication for its design (how complex the visual should be and how prominent the key messages should be). A research participant in Kennedy's research explains:

> *"Because I don't have a lot of time to read things […] if it's kept simple and easy to read, then I'm more likely to be interested in it and reading it all and […] to have a good look at it"* (J.C., 24, agricultural worker/engineer)

Considering the time an audience can dedicate to a visualisation is particularly important when designing a visualisation for policy audiences. As discussed in section 2.12, policymakers are time-poor, putting a premium on research outputs that do not require large time investments on the readers' part.

As well as skills and time, audience capacity extends to external factors including access to technology and bandwidth availability (Global Health Learning Center 2015). Elaborate interactive data visualisations may have the potential to engage the audience and promote research uptake, but if technological limitations prevent access and use, the potential contribution of the data visualisation to the objectives of the publisher will not be realised. Shirona Patel, from South Africa, notes:

> *"In SA, more and more people have access to cell phones and smartphones. However, most websites are not built to be responsive so often the use of tables, graphs and data is discouraged as they 'fall off the pages' from websites that do not adapt to mobile phones"*

Decisions regarding the production of data visualisation must therefore be guided by the audience's capacity to access, use and understand visualisations.

3.3 Teamwork and organisational buy-in

As well as audience capacity, data visualisation requires the producer to have a certain capacity. In particular, teamwork and organisational buy-in are important in the production of data visualisation.

Finally, decisions regarding investments in data visualisation are influenced by the degree of organisational buy-in. The production of data visualisation – particularly more-complex, interactive examples – requires an array of different skillsets, many of which are not traditionally associated with research communication. Among the skills required are data analysis (to clean data and draw out trends and relationships within the data), visual design (to create a

visually appealing product), digital skills (to create the product) and storytelling or journalism skills (to explain the significance and implications of the data) (Knezovich 2015).

Skills required for data visualisation production
· Communication/storytelling skills
· Digital/technological skills
· Data analysis/research skills
· Visual design skills

Identifying staff with these skillsets can be a challenging and a costly endeavour, as research participants explain:

> *"Having a team dedicated to data journalism is not cheap. Journalists must be hired or trained with statistical and mathematical analysis, it is not easy to find those profiles and if you are lucky [to find them] they must be paid for these skills"* (Hassel Fallas, Costa Rica)

> *"If an organization cannot bring in experts from outside, the onus is on the officials within the organization. And not everyone can produce brilliant [visualisations] – let alone PowerPoint presentations"* (Lebo Majara, South Africa)

Branching into the production of data visualisation requires an element of experimentation, learning-by-doing and investments in building staff capacity. To produce data visualisation therefore requires organisational buy-in.

> *"I've trained many journalists, communicators and researchers on digital storytelling, including data viz, but I find the biggest barrier is buy-in from management, and thus time and resource allocation for creating effective data visualisations, data journalism projects and other digital stories in-house remains limited"* (Anina Mumm, South Africa)

As well as investing in training and making the space for experimentation, the production of data visualisation also often requires staff with different skillsets to work together on the production. Ricci Coughlan, senior designer at the UK Department for International Development's Creative Content Team, explains:

> *"Having data scientists and statisticians who are data visualisation literate is a big challenge, as is finding graphic designers who are data literate. Creating high-quality visualisations tends to be the result of great teamwork between a number of different skillsets, no one individual can be expected to do everything … I think many organisations will likely already have these skills (designers, researchers, writers, statisticians), but they may have just not got them working as a team before on such a product. They may all be used to contributing their skills to a report but less experienced in working with visualisations together"*

Teamwork is essential not only to produce all of the elements of a visualisation (e.g. data analysis, digital design, storytelling), but also to ensure that the story being told by the visualisation is true to the data set:

> *"Sometimes advanced data analysis skills are needed to get data into a format that is useful, whether that means cleaning or finding trends and significant numbers to report on. In terms of visualising complex scientific data, it is essential to partner with the researcher and*

perhaps one independent researcher to ensure that the information is still presented correctly and completely in the final data viz" (Anina Mumm, South Africa)

Teamwork is therefore essential to ensure that the data visualisation does not mislead or misinform the reader. If the integrity of the visualisation is compromised through poor data analysis, poor design or poor storytelling, its potential to contribute to research communication goals, discussed in section 2, will not be realised.

3.4 Summary of the requirements for data visualisation production

This section has discussed the necessary conditions for data visualisation to be an appropriate form of research communication. Firstly, the availability of high-quality, interesting data was identified as a precondition producing data visualisations. Audience capacity to access and understand data visualisation was also discussed as a requirement, as was the need for the availability of bandwidth and technology among the target audience to be consistent with the design of the visualisation. Finally, production skills, teamwork and organisational buy-in were identified as necessary conditions on the production side to ensure that data visualisations have integrity and do not mislead or misinform readers through poor design or poor data analysis (cf. Burn-Murdoch 2013).

4. How can data visualisations be used to the greatest effect?

Where data visualisation is deemed an appropriate and valuable form of research communication, a number of factors can help to maximise the potential of the visualisation to contribute to the research communication goals discussed in section 2. These factors include timing and subject

---

Caveat: the need for empirical testing

What follows is a series of hypotheses based on the experiences of research participants. These hypotheses have not been empirically tested. More user testing is needed to identify a full set of factors that influence the effectiveness of data visualisations and help producers create visualisations that have the greatest impact.

---

relevance, contextualisation and knowledge of audience, and design integrity.

4.1 Timing and subject relevance

> Data visualisations are most effective when they are based on topical issues.

The effectiveness of a data visualisation, like other forms of communication, is influenced by the timeliness of its publication. Where the subject matter is topical and current, the data visualisation has a greater likelihood of having impact (Kwapien 2015; Young 2015). Amr Elsawy, an infographic designer in Egypt, explains:

> *"The story of the infographic should be one of the recent issues which matters to the public opinion"*

To maximise potential impact, Caleb Gichuhi of The Institute for Social Accountability in Kenya notes the importance of publishing data visualisations to coincide with 'policy windows', major conferences or public campaigns. Where data visualisations address unpopular issues, Gichuhi highlights the importance of creating a 'buzz' around the topic, for example by lobbying media outlets on the importance of the issue, before publishing the visualisation. Gichuhi uses the example of a visualisation on parliamentary bills to explain how this is achieved and to what effect:

"When we were about to release a four-part infographic that highlighted four bills in parliament that were showing how selfish and greedy the parliamentarians are, we began asking people questions like … do you know how greedy your member of parliament (MP) is? Do MPs promote human rights in their laws? Did you know MPs want to increase their salaries AGAIN? Then we would have a call to action … e.g. join our tweetchat tomorrow and answer the following questions and learn more … then tag as many actors, activists, politicians etc. This creates a buzz because people are eager to engage and share opinions and just when people are on social media for the tweetchat we 'launch' the infographs and then you have this engagement on social media about the infograph"

The importance of subject matter and its relevance was also identified in the Seeing Data project, which assessed the factors that influence readers' engagement with visualisations. Andy Kirk (2016) – a researcher on the project – explains that visualisation practitioners must consider topic interest when designing visualisations. Where interest in the topic is low, visualisation practitioners need to think about ways to create appeal and persuade readers to engage with the visualisation.

4.2 Contextualisation and knowledge of audience

> Data visualisations should be tailored to the target audiences' interests and needs to be most effective.

While timing and subject relevance are, to some extent, dependent on extraneous factors, effectiveness of a data visualisation can also be enhanced by contextualising data and, in doing so, enhancing the meaningfulness and utility of the visualisation. Contextualisation is imperative to demonstrate to the reader why the data matters and guide the user towards research uptake (the application of data to policy and practice) (Murray 2013; Thumer 2016).

To contextualise data requires the visualisation to be embedded within data journalism more broadly: that the story of the data is told whether through a well-designed and annotated visualisation or through supporting text (Diakopoulos 2013). Hassel Fallas from Costa Rica explains:

"The data viz tells us what is happening and journalists write articles that explain why such data is important. That relationship is what I think we should always consider when we work with data … Data journalism must emphasize the human side of the story. We must go out to study how the data behaves in the real world. Getting the best stories from people is mandatory to understand the context behind those data sets you are giving your faith. Without it, your story would be a cold and empty echo. Data journalism is for people and needs to be useful as well as meaningful."

The role of storytelling in contextualising data and drawing out its wider implications was highlighted through SciDev.Net and INASP's joint data challenge. The competition asked for researchers to 'pitch' their data. One of the winning data sets, which profiles a solar water heater installation at the Tenzinling Hotel in Paro, Bhutan, provided comprehensive data on various dimensions of the solar water heater (such as average incoming and outgoing water temperature per hour and average cloud cover). But a stand-alone visualisation of that data set was of limited relevance and applicability to SciDev.Net's audience. By contextualising the data set within Bhutan's changing energy markets, the story of the data set (which described solar energy's potential to replace hydroelectricity as the country's main export) was contextualised. The resulting data visualisation was more meaningful and more useful to a larger audience than an uncontextualised visualisation of the data set would have been.

Contextualising data therefore plays an important role in enhancing the effectiveness of visualisations, but doing so requires a clear understanding of the target audience (Global Health Learning Center 2015). Understanding the audience and designing visualisations that speak to their interests and requirements will enhance the likelihood that the visualisation will be read, engaged with and used.

Tailoring the story of the data visualisation to the audience's interests and needs must, however, be balanced by a trueness to the data set. Telling the story of a data set places parameters on storytelling, which are not necessarily in place in more traditional forms of journalism. Nicola Pearson, former joint editor of SciDev.Net, explains:

> "The data has to be the centre of the story. This means that the journalist has less flexibility in terms of storytelling. You can't just say what you want to say, instead you have to tell the story that is being told in the data"

Contextualising data and telling the story of its relevance and implications is a central component of an effective data visualisation, but requires that the story being told is tailored to the target audience while remaining true to the original data.

4.3 Design integrity

Design integrity also influences the effectiveness of a data visualisation. Those that follow good design principles and provide clear links to the original data are likely to be most effective.

As discussed in section 2.32, authoritativeness and trustworthiness are key factors in achieving research uptake and determining the effectiveness of a visualisation. Authoritativeness and trustworthiness are determined by a number of factors, including the source of the data and the integrity of the visualisation's design (Global Health Learning Center 2015). The effectiveness of data visualisation can therefore be enhanced by ensuring that data visualisations are well designed and follow standard guidelines of best practice in data visualisation production (see for example Kirk 2012; Lima 2009; Global Health Learning Center 2015). Hassel Fallas explains:

> "In [data visualisation,] we do not [do] storytelling as if it is putting together dense bricks of numbers and information. No, we are information architects and, as good architects, we align the logic and the accuracy of data analysis in harmony with good content, writing and appropriate visual display"

Research participants point out that a growing body of online users 'police' data visualisations, highlighting inaccurate or misleading examples. For example, Flowing Data – a website that provides tutorials and guides for data visualisation production has a Mistaken Data section that profiles visualisations that are poorly designed or based on misleading data. For example, they highlight visualisations which over-emphasise the difference between groups by using truncated Y axes (i.e. a vertical axis that does not start at 0) (see this blog post for example)

Visualisations that do not follow good practice risk misleading readers, resulting in a product that does not assist the formation of better informed decision-making processes. To minimise the potentially negative effects of data visualisation, the wider public must become better at critically appraising visualisations, while data visualisation producers must adhere to good practice principles. These principles relate not only to design but also to transparency around the data on which

visualisations are based ([Burn-Murdoch 2013](#)). The importance of citing the original source of data and providing a link to the full data set was emphasised by participants throughout this research as a particularly important principle for enhancing the trustworthiness of the visualisations. Caleb Gichuhi (Kenya) explains:

> *"It's prudent to share the source of the data"*

This sentiment is echoed by Lulu Pinney, freelance infographic designer and author of [Telling Information](#), who notes that providing a link to the full report or data set enables the reader to assess the credibility of the data for themselves (see also [Lima 2009](#)).

As well as enabling readers to assess the quality of data for themselves, trustworthiness can be increased by publishing your visualisation on a reputable platform, for readers often judge the quality and reliability of data based on the reputation of the publishing outlet and supporting foundations ([Kennedy 2015](#); DFID forthcoming). Clearly citing reputable associated or supporting organisations such as international donors can also be used as a proxy for trustworthiness, increasing the readers' confidence in the reliability of information provided.

4.4 Summary of ways in which to enhance the effectiveness of data visualisation

This section has discussed ways in which data visualisation can be used to greatest effect. Firstly, timing and subject relevance were discussed as influencing factors in determining the impact of data visualisation, and the benefits of creating a 'buzz' around unfashionable issues prior to publishing a data visualisation was briefly touched on. Contextualising data within broader issues or concerns was also discussed as a strategy for enhancing the effectiveness of data visualisations, and the corresponding need to know your audience was explored. Finally, design integrity – particularly citing the source of the data – was discussed as an important way to enhance the effectiveness of a data visualisation.

5. Conclusions

Data visualisation is a rapidly growing area of investment and interest, both within the research communication sector and the broader media landscape. This report has explored how data visualisation can contribute to the goals of research communication. These contributions are potentially significant, particularly in terms of generating greater public demand for evidence-based decision-making and supporting decision-makers to access, understand and apply data to policy and practice.

While the potential benefits of data visualisation are significant, concentrated and informed discussions around what data visualisations can add to research communication efforts and how these contributions can be achieved has been lacking. This report has attempted to address this gap by putting forward a set of hypotheses regarding the benefits of data visualisation and the appropriate use of such visualisations.

These hypotheses require further empirical testing and scrutiny so researchers, research communicators and journalists can make more-informed decisions about investing in data visualisation production. Empirical data on the relative benefits of data visualisation – compared with other forms of communication such as text and multimedia – to different research communication goals is needed to enable more-informed decisions about when to invest in visualisations. Empirical data on simple webmetrics, such as unique page views, shares and likes, would go some way to understanding the benefits of data visualisation (particularly in terms of promoting the selection of research outputs from the pool of information available to online users), but this would require it being shared between organisations. Information on the extent to which

data visualisation can contribute to research uptake is more difficult to obtain and requires analysis of user behaviours. Investments in such studies could make important contributions to understanding the pathways to research uptake and the role that data visualisation can play in promoting greater application of evidence to policy and practice.

Appendix A: Methodology

This report draws on virtual focus group discussions, in-depth interviews and existing literature to explore the contributions of data visualisation to research communication goals. Participants' responses were analysed alongside existing literature to identify key themes relating to the objectives and outcomes of data visualisations within the sphere of research communication. Where possible, qualitative data has been supported with quantitative monitoring data collected through SciDev.Net's established monitoring and evaluation framework.

Interviewees and online discussants were selected through a snowball sampling method[4] and include data visualisation experts, researchers, knowledge intermediaries and journalists. A total of 23 participants from Sub-Saharan Africa, Asia, Latin America, the Middle East, North Africa and the United Kingdom participated in the research, reflecting a range of different viewpoints and experiences. For a full list of participants, please see Appendix B.

Three online focus group discussions were held, each lasting for seven days. For the focus group discussions, participants were split into three groups according to their area of expertise, with one group consisting of journalists and researchers, a second of data visualisation experts (designers, producers etc.), and a third for NGO and civil society organisation (CSO) practitioners. It was decided to split participants by profession rather than by geographic location so the questions around which the discussions were ordered could be tailored to each groups' expertise and degree of familiarity with data visualisations. The questions on which the focus group discussions were based can be found in Appendix C.

To manage for the geographical spread of participants and to enhance the cost-effectiveness of this research, focus group discussions were held virtually. A 'bulletin board' tool was used in which the research topic was broken down into different threads. Participants were invited to contribute to each of the threads, which were numbered to guide respondents through a logical pathway of questions.

The bulletin boards were open for seven days to maximise the response rate from participants amid their busy schedules. Forty four percent of those who were confirmed their interest in participating contributed in some way. Email prompts were sent to inactive members of the focus group and the rate of contributions increased once the first participant in each group provided answers to the questions posed in each thread.

The virtual nature of the focus groups and their duration provided respondents with time and space to reflect on the questions before providing answers. This space for thought was reflected in the quality and length of answers provided by respondents. However, the duration of the focus group discussions required significant time investment on the part of the facilitator and meant that follow-up questions could not always be posted immediately. Thus, there was a trade-off between practicality and immediacy, which meant that, while the research was conducted in a more cost-effective manner, the dialogue stimulated through the focus groups was less dynamic than in more traditional face-to-face focus group discussions.

Limitations of the research

While this research draws on the expertise of stakeholders from a range of different countries and a spread of different sectors, the sample size is small. Therefore, findings cannot be said to be representative of any country or sector. Identifying participants proved a challenge in some regions

---

[4] Snowballing sampling is a method by which existing participants identify additional participants from their networks

and sectors. SciDev.Net's regional staff were asked to suggest potential participants but, due to limited exposure to or infiltration of data visualisations in some regions, particularly West Africa, few suggestions could be made in some instances. Similarly, identifying participants in the NGO/CSO sector with experience in using data visualisations was particularly challenging, again indicating that data visualisation is not a mainstream communication tool for some stakeholders.

Another limitation of this report is the lack of empirical evidence on how readers react to and use data visualisations. This lack of data means that the hypotheses put forward by research participants cannot be empirically verified, though they were supported by quantitative data where possible.

Appendix B: Research participants

Name	Country of residence	Job title	Organisation
Amr Elsawy	Egypt	Infographic designer	Al-Ahram Publishing House
Andrew Lee	United Kingdom	Head of digital	SciDev.Net
Anina Mumm	South Africa	Science communication and digital media specialist	ScienceLink SciBraai
Caleb Gichuhi	Kenya	Communications officer	The Institute for Social Accountability
Dr David Girling	United Kingdom	Director of research communication	School of International Development, University of East Anglia
Dr David Tarrant	United Kingdom	Data science trainer	The Open Data Institute
Dr Fatou Gueye	Senegal	Academic support officer	USAID Education and Research in Agriculture project
Gilbert Nakweya	Kenya	Freelance journalist	Freelance
Hassel Fallas	Costa Rica	Data-driven editor	La Nacion
Lebo Majara	South Africa	Science communicator	South African National Biodiversity Institute
Lulu Pinney	United Kingdom	Freelance infographic designer	Telling Information
Mohamed Ali Eddin	Egypt	Photo-journalist	Freelance
Naveen Gattu	India	Co-founder	Gramener
Nicola Pearson	United Kingdom	Former assistant editor	SciDev.Net
Nigel Hawtin	United Kingdom	Information graphics and visual designer	Nigelhawtin.com
Prachi Salve	India	Senior policy analyst	IndiaSpend
Ricci Coughlan	United Kingdom	Senior designer	Creative Content Team, UK Department for International Development
Shirona Patel	South Africa	Head of communications	University of Witwatersrand
Srikanth Viswanathan	India	Coordinator	Janaagraha Centre for Citizenship and Democracy
Surendran Balachandran	India	Growth and partnerships	SocialCops
Dr Tom Smith	United Kingdom	Director	Oxford Consultants for Social Inclusion
Yalaoui Bilal	Algeria	Researcher in applied mathematics and information science	Freelance
Dr Zipporah Ali	Kenya	Executive director	Kenya Hospices and Palliative Care Association

Appendix C: Focus group questions

Questions asked to all participants:
- What is the purpose of data visualisation?
- What are the main strengths/benefits of data visualisation?
- What are the main challenges associated with data visualisations?
- How can data visualisations be made more effective in informing decision-making processes?

Additional questions asked to data visualisation experts:
- What makes a 'good' visualisation? What makes a 'bad' visualisation?

Additional questions asked to journalists:
- What is your experience with data visualisation? How common are they in your country?

Additional questions asked to development practitioners:
- What is your experience with data visualisation? How common are they in your country?
- What strategies does your organisation use to promote its cause or values?

Semi-structured interviews were based on this set of questions but allowed scope for the discussion to evolve organically.