**Raising the Bar Graph: Upgrading to Interactive Dynamic Data Visualizations in Psychology**

“Via dynamic and interactive graphics, today’s technology allows students to visualize externally what they have difficulty representing mentally”

- *David Moreau (2015), p. 2*.

Research on human perception have taught us a lot about how to better communicate data through illustrations and graphics. Visualizing data-driven results continues to improve through psychological research and technological innovations. Today, researchers have greater capacity to collect better-quality data and in larger quantities (e.g., online surveys, cell phone applications, wearable technologies, social media, Google Analytics, etc.). In order to preserve the integrity and purpose of our graphs, data visualizations must be able to accommodate this surge in information volume and complexity of the digital age. Yet, despite the exponential growth in computing power and software tools, psychology researchers continue to rely mostly on static data visualizations.

**The (Optical) Illusion of Current Data Visualization Practices**

Static visualizations are entirely acceptable and appropriate for many purposes (as I hope the figures in this paper), however, as we transition into the big data era, relying *only* on static graphics may not be enough (Heer & Kandel, 2012). Static graphs can limit the type and amount of information one wishes to communicate, or the speed and comprehension of details the reader is meant to perceive; certain types of static graphs can often be no more than “visual tables” (Weissgerber et al., 2015, p. 1). In a 2014 systematic review of research articles published in the top 25% physiology journals, the authors found that the most common type of data visualizations are static graphs. And specifically, static *bar graphs* (Weissgerber et al., 2015). Although a *recent*[[1]](#footnote-1) systematic review in psychological research has yet to be published (to my knowledge), it appears that psychology suffers from the same graph use issues as physiology and other disciplines. The #barbarplots project (2017) examined 131 research articles from the first six month of 2016 in four high impact psychology journals and found that there is a substantial presence of static bar graphs. Particularly, from a total of 104 graphs presented, 55% of figures were static bar graphs.

Static bar graphs have been heavily criticized (especially with continuous data) for increasing the risk of misinterpreting research findings and providing limited information (Cooper et al., 2002; Gelman, 2017; Lane & Sándor, 2009; Schriger et al., 2006; Saxon, 2015; Weissgerber et al., 2015). For example, a static bar graph can portray potentially countless different distributions (e.g., bimodal, skewed, containing outliers, difference in sample size) the same way if their means and variability are similar (see Figure 1).

Chart, scatter chart

Description automatically generated

**Figure 1.** **Many different datasets can lead to the same static bar graph.** The full data may suggest different conclusions from the summary statistics. The means and standard errors (SEs) for the four example datasets shown in Panels B–E are all within 0.5 units of the means and SEs shown in the bar graph (Panel A). In Panel B, the distribution in both groups appears symmetric. Although the data suggest a small difference between groups, there is substantial overlap between groups. In Panel C, the apparent difference between groups is driven by an outlier. Panel D suggests a possible bimodal distribution. Additional data are needed to confirm that the distribution is bimodal and to determine whether this effect is explained by a covariate. In Panel E, the smaller range of values in group two may simply be because there are only three observations. Additional data for group two would be needed to determine whether the groups are different. Figure and caption adapted from Weissgerber et al. (2015).

This can mislead readers to make a certain set of assumptions (e.g., normality, equal sample size and variability, no outliers, etc.) which in reality might have never been met. In addition, if non-independent – or paired – data are presented using static bar graphs, readers might falsely infer independence and ignore within-subject differences (or lack thereof; see Figure 2).

Diagram, engineering drawing

Description automatically generated

**Figure 2.** **Additional problems with using static bar graphs to show paired data.** The bar graph (mean ± SE) suggests that the groups are independent and provides no information about whether changes are consistent across individuals (Panel A). The scatterplots shown in the Panels B–D clearly demonstrate that the data are paired. Each scatterplot reveals very different patterns of change, even though the means and SEs differ by less than 0.3 units. The lower scatterplots showing the differences between measurements allow readers to quickly assess the direction, magnitude, and distribution of the changes. The solid lines show the median difference. In Panel B, values for every subject are higher in the second condition. In Panel C, there are no consistent differences between the two conditions. Panel D suggests that there may be distinct subgroups of “responders” and “non-responders.” Figure and caption taken from Weissgerber et al. (2015).

Of course, there are numerous types of static data visualizations, other than bar graphs, that might be better suited for the researcher or data scientist’s needs (e.g., scatterplots, line graphs, violin plots). Yet, static graphs (unintentionally) force the reader to look at the data through a single lens – only the one that the authors intended. The authors have complete and utter control over the type of graph, which variables go in, the scaling or truncation of axes, the angle or point of view (e.g., in a 3D plot), the colours, sizes, and more. Many static graphs can also be inaccessible to some readers. For example, readers who are colour blind or have difficulty with depth perception, graphs that include very small or too many labels, etc.

Furthermore, static figures prevent readers from exploring the visualized data independently and assessing data-driven conclusions impartially: readers cannot manipulate the graph, inspect particular observations, rescale the axes, or visualize different analyses. By doing so, we unintentionally limit the information presented and therefore transparency. This might sound normal; it is the default after all, we have been seeing static figures in academic journals for decades. But we now have the capacity to allow readers to share control over how the data is visualized, so why do we not make it the new “normal?”

This unintentional lack of transparency is problematic in data visualization as it is in other research aspects such as reporting practices (e.g., limited/ambiguous reporting, excluding analyses due to lack of statistical significance, etc.). Reporting practices, among other research facets, have been directly addressed and are improving thanks to various taskforces, the open science framework (OSF), and more. Given recent concerns about the replication crisis, publication bias, and lack of transparency in psychological research, a reform in data visualization should also be considered.

**Enter Beyond-Static Graphics**

Interactive and dynamic data visualizations can accommodate virtually endless amount and types of information. In fact, with interactive dynamic visualizations, readers can tailor and modify the graphic to display the details of their choosing. But, before we go deeper into why we should implement these data visualization methods, we first need to define what are interactive and dynamic visualizations. In this paper, I define three types of non-static data visualizations: interactive, dynamic, and interactive dynamic data visualizations. However, I will mostly focus on the latter type, interactive dynamic, due its advanced features and benefits.

In an *interactive data visualization* (IDV) the user[[2]](#footnote-2) can probe the presentation to view certain aspects of the data (e.g., value on another variable), and thus *interact* with the display directly instead of using menus or cross-referencing a table (Ward et al., 2015). Other interactive features may include scaling or zooming in and out on a particular slice of data, highlighting certain categories or values, rotating (in 3D graphics) or panning over, and more. An example of IDV is presented in Figure 3 and can be accessed [here](https://plotly.com/r/3d-scatter-plots/).

Chart, scatter chart

Description automatically generated

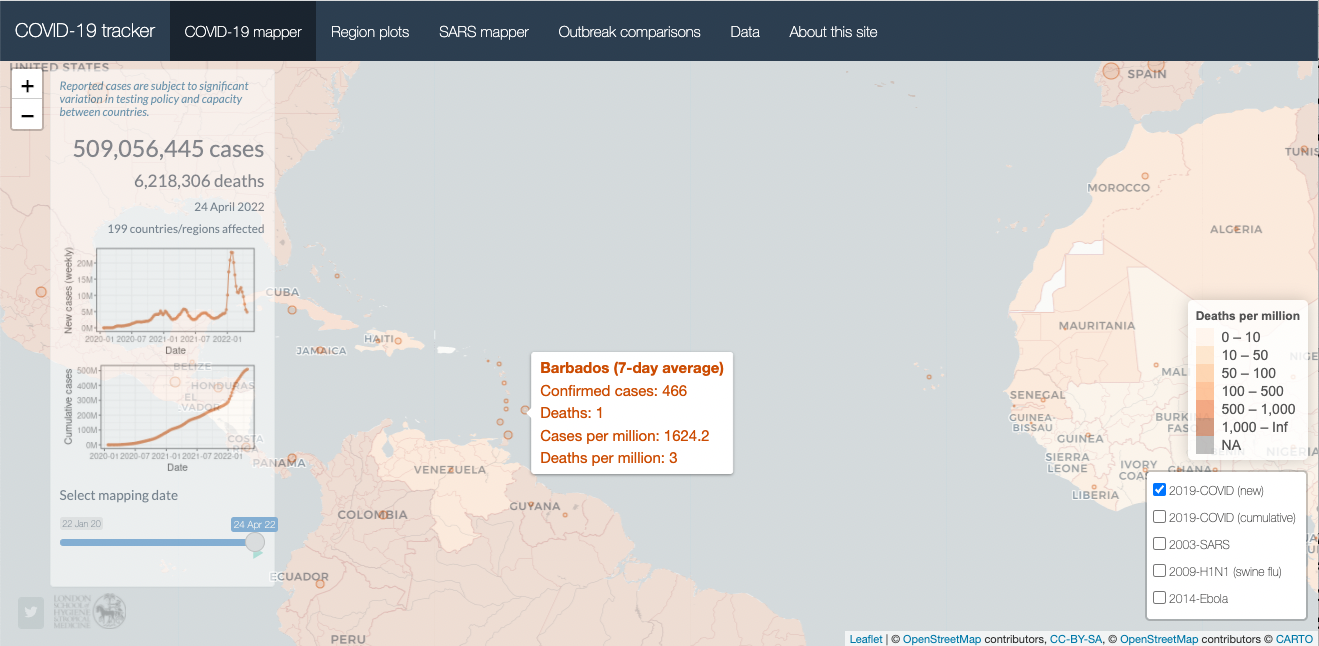
**Figure 3. Interactive data visualization.** This is a 3D scatter plot with interactive features. The dataset presented is ‘mtcars’ from the ‘datasets’ R package (R Core Team, 2022). The plot displays the relationship between different vehicle types (i.e., automatic in red and manual in blue), weight (x axis), time it takes to drive a quarter of a mile in minutes (y axis), and horsepower (z axis). The user can rotate the plot 360°, zoom in and out, pan the “camera” over, highlight one or more categories (e.g., red; automatic), hover over any individual observations to see its coordinates, and download a snapshot of the current position of the plot.

*Dynamic data visualizations* (DDV) are representations of data that can change their graphical make-up while presented (Ploetzner & Lowe, 2004; Schnotz et al., 1999). Kaput (1992) considered time as a dimension in DDV. This is, of course, true when visualizing time-series, hierarchical, or longitudinal designs, or to depict a process (e.g., neuron firing). However, in other designs, I argue, any variable – whether categorical or continuous – can serve as a dimension in DDV (e.g., treatment condition, IQ, etc.). In DDV, displayed values of the outcome can change (while still presented) as a function of another variable (e.g., time, group). In a linear, or automatic DDV, changes transpire on their own and cannot be altered by the user or reader. An example of DDV is presented in Figure 4 and is available [here](https://plotly.com/ggplot2/animations/#:~:text=Play-,Mulitple%20Trace%20Animations,-library(plotly)%0Alibrary).

Chart, scatter chart

Description automatically generated

**Figure 4. Dynamic data visualization.** This is a scatter plot with dynamic features. The dataset presented comes from the ‘gapminder’ R package (Bryan, 2017). The plot displays the relationship between gross domestic product per capita (gdpPercap; x axis) and life expectancy in years (lifeExp; y axis) across different times, from 1952 to 2007 (bottom scale). The user can zoom in and out, highlight one or more categories (e.g., blue; Europe), hover over any individual observations to see specific details (as seen in the red box), and when the “play” button is pressed (bottom left corner), changes across the years can be seen via animation.

Lastly, as its name suggests, *interactive dynamic data visualizations* (IDDV) combine elements from both IDV and DDV and is therefore the best of both worlds. In IDDV, any change in the graphical structure follows a direct action from the user. This way, users have control over – or, can manipulate – how the data is displayed (Ploetzner & Lowe, 2004). In many cases, the viewer even has control over how the data is analyzed (see below example by Ellis & Merdian, 2015). Schwan and Riempp (2004) noted that IDDV “enable the user to adapt the presentation to [their] individual cognitive needs” (p. 296). An example of IDDV is presented in Figure 5 and is available [here](https://haozhu233.shinyapps.io/shinyMRI-contest/).

**Figure 5. Interactive dynamic data visualization.** This app allows users to visualize and analyze COVID19 and SARS geographic and descriptive data across time using both interactive and dynamic features created by Parker (2022). Users can browse the different tabs at the top, adjust the date on the left blue scale, hover over any region or observation, apply filters/masks, pan over, zoom, view the raw data, and create a customized line graph.

Although there are mixed conclusions about the effectiveness of IDV and DDV (e.g., Hegarty, 2004; Hood et al., 2019; Ploetzner & Lowe, 2004; Rolfes et al., 2020; Schnotz et al., 1999; Schwan & Riempp, 2004), there is a growing body of literature advocating for the use of IDDV in both education and research settings (e.g., Ellis & Merdian, 2015; Heer, & Shneiderman, 2012; Rolfes et al., 2020; Ward et al., 2015).

**Why We Should Use Interactive Dynamic Graphics in Psychology**

There are many clear advantages to incorporating IDDV in academic articles and teaching. In this paper, I focus only on a few key reasons relating to open research practices, live data updates after visualization deployment, and effective data communication:

***IDDV Incite Transparent Research Practices***

Concerns about the credibility of previously published results in psychological research inspired new standards of recommended practices such as preregistration, registered reports, open data, shared analyses, etc. IDDV are much more aligned with the new standards of transparent research than static graphs. Thoughtful IDDV allow careful inspection of every dimension of the data, whereas with static visualizations, this is usually impossible or can make the graph too complex to comprehend. More advanced IDDV can even grant the reader the ability to explore and reanalyze the data on their own. This is perhaps one of the highest levels of open research practices. For example, in the [IDDV example created by Ellis and Merdian (2015)](https://psychology.shinyapps.io/example3/) using R Shiny app, the user can select the variables of their choice, differentiate by group, apply different filters or masks, and recalculate the regression equation and correlation coefficient (see Figure 6).

Graphical user interface, chart

Description automatically generated

**Figure 6. Interactive dynamic example with exploratory data analysis applications.** This interactive dynamic figure allows the user to select the variables of their choice (from the dropdown menus at the top left), differentiate by group (male, female), apply different filters or masks (previous victim, not previous victim), and recalculate the regression equation and correlation coefficient (bottom right table). The user can also inspect the data using a different graph type (e.g., boxplot) by switching to the Boxplot tab above the figure.

Furthermore, with tools such as R Shiny, where the data and analyses are stored on the author’s server, users can have all the benefits of having access to the full dataset without compromising participant confidentiality. One of the main factors preventing researchers from making their data openly available to external evaluators is the risk of breaching ethical protocols and exposing private information. However, using IDDV, the data can still be investigated and reanalyzed by the user *without actually having to share the data*. Note that we could explore and visualize data in the Ellis and Merdian (2015) example above without having access to the Fear of Crime dataset.

***Live Data Visualizations***

IDDV can continue to receive new data and update its visual structure accordingly, even after its deployment. Normally, once a static plot or graph are produced, the figure cannot be changed. This can influence our conclusions, too; once a paper is finalized, the figures are produced, scientists publish their paper and (usually) never look back. However, in the current era of big data and cloud technology, it is possible and often useful to stream and visualize our data live as it is being collected. Imagine you are conducting a study looking at various transactions on a social media platform (e.g., Facebook, Twitter, Instagram). Using IDDV, researchers can follow and visualize their data in real-time as social transactions take place. For example, there is an existing dashboard visualizing a livestream [downloads of R packages from CRAN](https://jjallaire.shinyapps.io/shiny-crandash/) (Sievert et al., 2022; see Figure 7). A [similar app](https://hadley.shinyapps.io/cran-downloads/) was attempted by Hadley Wickham (2019). This novel approach opens a lot of investigative avenues and provides useful tools to expend the arsenal of data visualization and research capabilities.

***Graphical user interface, website

Description automatically generated***

**Figure 7. Interactive dynamic Live Data Visualization.** livestream of download logs from cran.rstudio.com. This Shiny app illustrates the names and frequencies of R packages downloaded by users. Viewers can also manipulate the rate and capacity of the visualization, scan the raw logs in the top tab, review the source code, and share this app on various social media platforms.

***IDDV are More Effective at Communicating Data***

As Moreau beautifully put it, “[v]ia dynamic and interactive graphics, today’s technology allows students to visualize externally what they have difficulty representing mentally” (2015, p. 2). In a recent study, Hood and colleagues (2020) found that IDDV figures in digital publications were more effective at communicating main effects and null relationships than static graphs. In another study by Rolfes et al. (2020), students who were given materials visualized using IDDV performed significantly better than students in the static visualization group. These results suggest that IDDV might be more effective in communicating data. Furthermore, IDDV make active exploration possible where readers are free to interact with their data rather than trying to absorb it passively; functional engagement improves learning, interest, and comprehension (Bodemer et al., 2004). This is particularly helpful because direct interaction with visual content facilitates the involvement of the motor system (Wraga et al., 2003), which is essential for achieving deeper levels of understanding (Moreau, 2015).

**Concluding Remarks**

Designing and creating IDDV used to be a software developer’s job and required extensive programming skills in Java, HTML, or CSS. This certainly might have been an obstacle and deterrent for many. Fortunately, however, this is no longer the case. Nowadays, scientists and authors with working knowledge of R (R Core Team, 2022) can generate their own *free* IDDV using packages such as Shiny (RStudio Team, 2022), flexdashboard (Sievert et al., 2022) and esquisse (Meyer & Perrier, 2022) relatively easily. Of course, making IDDV is also possible without programming at all when using Tableau, Figma, and more. Though, these platforms might include fees.

IDDV also make graphs and figures much more accessible. IDDV allow users without prior software knowledge or fee-dependent license to explore and analyze data. The user-friendly interface implemented in most IDDV does not require programming abilities or downloading/purchasing software, which are otherwise required to manipulate and analyze data. In addition, shifting the control of various visual features (e.g., colour, position, size, speed) to the user can make graphs more convenient. Each user can tailor the graph to their own individual needs. To that extent, IDDV can accommodate wider range of abilities (e.g., colour blind, short sightedness, slow/quick processing speed, etc.).

That said, it is important to note that IDDV are not a silver bullet. Incorporating interactive and dynamic features to your graphs carelessly will not make them better. Designing your IDDV requires a lot of thought, planning, and work, probably even more than static graphs. For this reason, as with static graphics, it is useful for individuals to follow recommend practices and familiarize themselves with taxonomy and techniques (e.g., Heer & Schneiderman, 2012).

Incorporating IDDV into mainstream academia will not be easy, it certainly will require a long adjustment period and a steep learning curve (pun not intended). However, continuing to integrate new technologies into the field of data visualization is inevitable: it will happen in academia as it is happening in industry where large corporations have already begun applying the “fourth paradigm” (Tansley & Tolle, 2009) of advanced computing and (big) data-driven scientific breakthroughs.

Finally, the examples, reasons, and topics in this paper only begin to scratch the surface. Dynamic and interactive data visualizations are not entirely new. In fact, one occurrence of an interactive system for visualizing probability plots dates back to the late 1960s (Fowlkes, 1969, as cited in Friendly, 2008). Still, there is much more to learn about how interactive and dynamic features affect our perception and how we can harness these tools to improve data communication.

**References**

Bar Bar Plots Project (2017). #barbarplots. Available online at: <https://www.kickstarter.com/projects/1474588473/barbarplots>. (Accessed 2022-04-20).

Best, L. A., Smith, L. D., & Stubbs, D. A. (2001). Graph use in psychology and other sciences. *Behavioural processes, 54*(1-3), 155–165. <https://doi.org/10.1016/s0376-6357(01)00156-5>

Bodemer, D., Ploetzner, R., Feuerlein, I., & Spada, H. (2004). The active integration of information during learning with dynamic and interactive visualisations. *Learning and instruction, 14*(3), 325-341. <https://doi.org/10.1016/j.learninstruc.2004.06.006>

Bryan, J. (2017). Data from Gapminder. <https://github.com/jennybc/gapminder>

Cleveland, W. S. (1984). Graphs in Scientific Publications. *The American Statistician, 38*(4), 261–269. <https://doi.org/10.2307/2683400>

Cooper, R. J., Schriger, D. L., & Close, R. J. (2002). Graphical literacy: the quality of graphs in a large-circulation journal. *Annals of emergency medicine, 40*(3), 317–322. <https://doi.org/10.1067/mem.2002.127327>

Ellis, D. A., & Merdian, H. L. (2015). Thinking outside the box: developing dynamic data visualizations for psychology with Shiny. *Frontiers in psychology, 6*, 1782. <https://doi.org/10.3389/fpsyg.2015.01782>

Fowlkes, E. B. (1969). *User’s manual for a system for interactive probability plotting on Graphic-2*. Tech. rep., Bell Laboratories.

Friendly, M. (2008). *A Brief History of Data Visualization. In: Handbook of Data Visualization*. Springer Handbooks Comp. Statistics. Springer, Berlin, Heidelberg. <https://doi.org/10.1007/978-3-540-33037-0_2>

Gelman, A. (2017). *Graphs as comparisons: a case study*. Blog Post. Available online at: <http://andrewgelman.com/2017/07/16/graphs-comparisons-case-study>

Heer, & Kandel, S. (2012). Interactive analysis of big data. *Crossroads (Association for Computing Machinery), 19*(1), 50–54. <https://doi.org/10.1145/2331042.2331058>

Heer, & Shneiderman, B. (2012). Interactive dynamics for visual analysis. *Communications of the ACM, 55*(4), 45–54. <https://doi.org/10.1145/2133806.2133821>

Hegarty. (2004). Dynamic visualizations and learning: getting to the difficult questions. *Learning and Instruction, 14*(3), 343–351. <https://doi.org/10.1016/j.learninstruc.2004.06.007>

Hood, Graber, C., & Brase, G. L. (2019). Comparing the Efficacy of Static and Dynamic Graph Types in Communicating Complex Statistical Relationships. *Frontiers in Psychology, 10*, 2986–2986. <https://doi.org/10.3389/fpsyg.2019.02986>

Kaput, J. (1992). *Technology and mathematics education*. Handbook of research on mathematics teaching and learning, 515, 556. <https://doi.org/10.2307/749579>

Lane, D. M., & Sándor, A. (2009). Designing better graphs by including distributional information and integrating words, numbers, and images. *Psychological methods, 14*(3), 239–257. <https://doi.org/10.1037/a0016620>

Marmolejo-Ramos, & Matsunaga, M. (2009). Getting the most from your curves: Exploring and reporting data using informative graphical techniques. *Tutorials in Quantitative Methods for Psychology, 5*(2), 40–50. <https://doi.org/10.20982/tqmp.05.2.p040>

Meyer, F. & Perrier, V. (2022). esquisse: Explore and Visualize Your Data Interactively. https://dreamrs.github.io/esquisse/, <https://github.com/dreamRs/esquisse>.

Moreau. (2015). When seeing is learning: dynamic and interactive visualizations to teach statistical concepts. *Frontiers in Psychology, 6*, 342–342. <https://doi.org/10.3389/fpsyg.2015.00342>

Pastore, M., Lionetti, F., & Altoè, G. (2017). When One Shape Does Not Fit All: A Commentary Essay on the Use of Graphs in Psychological Research. *Frontiers in psychology, 8*, 1666. <https://doi.org/10.3389/fpsyg.2017.01666>

Parker, E. (2022). COVID-19 Tracker. <https://vac-lshtm.shinyapps.io/ncov_tracker/>

Ploetzner, R., & Lowe, R. (Eds.). (2004). Guest editorial: Dynamic visualisations and learning [Editorial]. *Learning and Instruction, 14*(3), 235–240. <https://doi.org/10.1016/j.learninstruc.2004.06.001>

R Core Team (2020). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. <https://www.R-project.org/>

Rolfes, T., Roth, J., & Schnotz, W. (2020). Learning the Concept of Function With Dynamic Visualizations. *Frontiers in psychology, 11*, 693. <https://doi.org/10.3389/fpsyg.2020.00693>

RStudio Team (2022). RStudio: Integrated Development for R. RStudio, Inc., Boston, MA. <http://www.rstudio.com/>

Saxon, E. (2015). Beyond bar charts. *BMC Biol. 13*:60. <https://doi.org/10.1186/s12915-015-0169-6>

Schnotz, W., Böckheler, J., & Grzondziel, H. (1999). Individual and co-operative learning with interactive animated pictures. European Journal of Psychology of Education, 14(2), 245–265. <https://doi.org/10.1007/BF03172968>

Schriger, D. L., Sinha, R., Schroter, S., Liu, P. Y., & Altman, D. G. (2006). From submission to publication: a retrospective review of the tables and figures in a cohort of randomized controlled trials submitted to the British Medical Journal. *Annals of emergency medicine, 48*(6). <https://doi.org/10.1016/j.annemergmed.2006.06.017>

Schwan, S., & Riempp, R. (2004). The cognitive benefits of interactive videos: Learning to tie nautical knots. *Learning and instruction, 14*(3), 293-305. <https://doi.org/10.1016/j.learninstruc.2004.06.005>

Sievert C, Iannone R, Allaire J, Borges B (2022). flexdashboard: R Markdown Format for Flexible Dashboards. <https://pkgs.rstudio.com/flexdashboard/>

Tansley, S., & Tolle, K. M. (2009). The fourth paradigm: data-intensive scientific discovery (Vol. 1). A. J. Hey (Ed.). Redmond, WA: Microsoft research.

Ward, M.O., Grinstein, G., & Keim, D. (2015). *Interactive Data Visualization: Foundations, Techniques, and Applications* (2nd ed.). A K Peters/CRC Press. <https://doi.org/10.1201/9780429108433>

Weissgerber, T. L., Milic, N. M., Winham, S. J., & Garovic, V. D. (2015). Beyond bar and line graphs: time for a new data presentation paradigm. *PLoS biology, 13*(4), e1002128. <https://doi.org/10.1371/journal.pbio.1002128>

Wickham, H. (2019). CRAN downloads. <https://hadley.shinyapps.io/cran-downloads/>

Wraga, M., Thompson, W. L., Alpert, N. M., & Kosslyn, S. M. (2003). Implicit transfer of motor strategies in mental rotation. *Brain and cognition, 52*(2), 135-143. <https://doi.org/10.1016/S0278-2626(03)00033-2>

Zhu, H., Chand, N. M., & Travison, T. (2019). Shiny MRI. Available at: <https://haozhu233.shinyapps.io/shinyMRI-contest/>

1. Some less recent systematic reviews exist, see Best et al. (2001) and Cleveland (1984). Results from these reviews do not differ dramatically from those of the #barbarplots project (2017) and Weissgerber et al. (2015). [↑](#footnote-ref-1)
2. I will use the terms user, reader, and viewer interchangeably from now on when referring to the “audience” or consumers of the graphics, as opposed to the researcher, author, or scientists who design/produce the graphics. [↑](#footnote-ref-2)