Moving Beyond the Bar Plot and the Line Graph to Create Informative and Attractive Graphics[1]

JENIFER LARSON–HALL
The University of Kitakyushu, Department of Foreign Languages, 4-2-1, Kitagata, Kokouraminami-ku, Kitakyushu, 802–8577, Japan Email: larsonhall@kitakyu-u.ac.jp

Graphics are often mistaken for a mere frill in the methodological arsenal of data analysis when in fact they can be one of the simplest and at the same time most powerful methods of communicating statistical information (Tufte, 2001). The first section of the article argues for the statistical necessity of graphs, echoing and amplifying similar calls from Hudson (2015) and Larson–Hall & Plonsky (2015). The second section presents a historical survey of graphical use over the entire history of language acquisition journals, representing 192 years of journal publishing. This shows that a consensus for using certain types of graphics, which lack data credibility, has developed in the applied linguistics field, namely the bar plot and the line graph. The final section of the article is devoted to presenting various types of graphic alternatives to these two consensus graphics. Suggested graphics are *data accountable* and present *all* of the data, as well as a summary structure; such graphics include the scatterplot, beeswarm, or pirate plot. Such graphics attract readers, help researchers improve the way they understand and analyze their data, and build credibility in numerical statistical analyses and the conclusions that are drawn.

Keywords: methodology; statistics; graphics; data; second language acquisition

GRAPHICS ARE TOOLS FOR HELPING US think about quantitative data in a way that most of us cannot do without visualization; at their best, they can make numbers and words come alive for the reader. Tufte (2001) stated that "of all methods for analyzing and communicating statistical information, well-designed data graphics are usually the simplest and at the same time the most powerful" (p. 9).

Certainly, a research conclusion like "the amount of vocabulary produced by bilingual children depends on the percentage of their input in each language" is not enough proof to convince other researchers that the conclusion is justified. However, many readers *are* satisfied with seeing a statistical number that represents the strength of the relationship, such as $r = .50$. By contrast, this article will argue that researchers should not be satisfied with only a number; instead, where

possible,[2] *data accountable* graphics should always accompany these types of numbers because such graphics will firmly establish the credibility of the statistical arguments. In other words, readers should not really feel convinced of a statistical conclusion until they have seen both numbers and *data accountable* or at least *data rich* graphics.

Here, the term *data accountable* graphics is defined as graphics that plot *all* of the relevant details of the dataset, such as, for example, a scatterplot, which plots the amount of vocabulary against the amount of input received by the bilingual child, or pirate plots for group data, which show distributions and summary data as well as the individual data points. *Data rich* graphics are defined as graphics that show the distribution of the data and necessarily present a large amount of information about the data set to the reader, although they do not show individual points. Some examples of this type of graphic are the histogram and the boxplot.

This article will first outline the argument that graphics are just as vital to statistical analysis as numbers and statistical tests. Second, it

will present evidence from a formal survey of graphics over the history of second language research. That review will show that, within the field, the wrong kinds of graphics are currently widely prevalent and that, historically, this area has seen regression instead of progression. The remainder of the article will present examples of informative graphics that can be fruitfully used with language research data.

WHY GRAPHICS ARE DESIRABLE

There are a number of convincing arguments for why graphics should be considered not merely ornamental flourishes to a sound statistical argument, but in fact are a necessary component of the same (Cleveland, 1985; Tufte, 1990; Tukey, 1977). Proponents have insisted that such graphics must present a full picture of the data, not just a summary—what I am here calling data accountable or data rich graphics.

As second language (L2) research has become more sophisticated, methodological improvements have been called for. For example, Norris, Ross, and Schoonen (2015), in a special methodologically oriented issue of *Language Learning*, recently called for improvement in the way "quantitative research is conceived, the ways in which data are collected, the analyses that are employed in making sense of patterns observed, and the typical approaches taken to reporting findings" (p. 5). In a similar vein, editors of major L2 publications (e.g., Byrnes, 2013; Ellis, 2015) have noted that a methodological revolution currently seems to be taking place. Not surprisingly, researchers may feel somewhat overwhelmed by the calls for methodological improvements. In response, this article highlights the use of graphics and argues that, as a tool, they are not very complicated while making the understanding of statistics much simpler. Among their obvious advantages are these: Graphics are inherently more attractive to many people than statistical numbers; they make finding patterns for diagnostic as well as data analysis functions simpler; and they reveal patterns that cannot otherwise be seen in summary numbers. That means they can lead to more discoveries in the data and more accurate data analysis, add weight to a statistical argument, and can even function as abstracts to provide a quick summary of the results of a study.

Graphics Attract Readers

Carefully chosen graphics can make research more accessible and attractive to readers. DeKeyser, former editor of *Language Learning* and current associate editor of *Bilingualism: Language and Cognition*, opined that graphs catch the most attention when people thumb through articles (personal communication, 2013). Byrnes, current editor of *The Modern Language Journal*, noted that most people who are not interested in methodological issues don't really understand statistics very well and are intimidated by them. By contrast, when used correctly, graphics can help readers judge more readily the statistical significance of a finding, not least because they tend to be less formidable than the at times overwhelming statistical jargon in an article.

*Visual Pattern Finding Is Almost Effortless*

Another good reason for using graphs is that making perceptual inferences from visual evidence comes easily to humans. Ware (2008) observed that the "activation of meaning from an image generally occurs in a small fraction of a second, much less time than it takes to read a paragraph of text. This activation through a single glance makes images far more efficient than words at conveying certain kinds of information" (p. 107). One kind of efficiency is that an extremely large amount of data can be digested very quickly. More specifically, Tufte (2001) claimed that the human eye can locate 100 points in one square centimeter, and that this is done automatically and implicitly. This makes understanding a graph quite easy, once the variables being graphed have been decoded. Another kind of efficiency is that the brain also judges geometrical aspects of a graphic, such as position, size, and slope of lines, at a mostly unconscious level. Humans are especially good at making comparisons from visuals. Larkin and Simon (1987) argued that graphics place all of the pertinent information together in one place, avoiding a search for the data needed for making inferences. In that way, they can provide simple yet powerful ways of seeing patterns in the data. It is, then, not difficult to agree with Ware (2008) who noted: "Often, to see a pattern is to understand the solution to a problem" (p. ix).

Graphics Reveal Patterns Not Seen in Data Summaries

Besides those general advantages, methodologically weightier arguments in support of the use of graphics exist as well. Cleveland (1985) pointed out that statistical methods for presenting data, such as giving a mean and standard deviation for a group, calculating a correlation coefficient or

FIGURE 1

Graphic From Weissgerber, Milic, Winham & Garovic (2015): Univariate Scatterplot Showing the
Inconclusiveness of Data Sets With the Same Mean and Standard Deviations
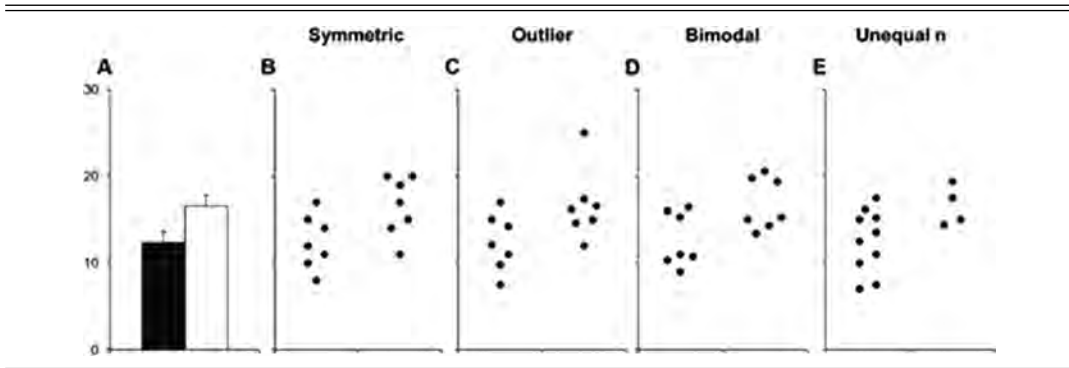
giving a *t*-test result, reduce the amount of data available. As a result, patterns in the data that could be easily uncovered by looking at a graphic are often hidden when only data summaries are provided. To Tukey (1977), "summaries can be very useful, but they are not the details …. there will be no substitute for having the full detail where we can look at it, set out in as clear a way as we can easily manage" (p. 27). For example, Figure 1 from Weissgerber et al. (2015) shows how much more information about the shape of the distribution is included in univariate scatterplots of the data as opposed to the graphed means and standard deviations. All of the distributions shown (symmetric, with outliers, bimodal or unequal *n*) are hidden when a barplot summarizing the mean score and standard deviation is shown. Indeed, conclusions about the differences between groups could be widely different for each of the four distributions shown in Figure 1. More important, this figure vividly illustrates the critique that a lack of distributional information can actually *hinder* scientific evaluation, as pointed out almost two decades ago by Wilkinson and the APA task force (Wilkinson, 1999). For example, in the panel labeled "Outlier" in Figure 1, the distributions are quite similar except for one outlier, so that the apparent difference seen in the barplot may be spurious; the bimodal nature of the data in the "Bimodal" panel means that there may be two separate patterns inside the existing groups, and a third factor that splits both groups should be examined. The "Unequal *n*" panel shows that it is difficult to say much about the two different groups since one group has so few data points, thereby highlighting the need to obtain more data for that group. At a deeper level, then, data accountable graphics will not only help readers of a research report ascertain the shape of the

data distribution; much earlier they will alert researchers to the need to explore their data sets before drawing any conclusions. (This point is discussed and illustrated in more detail in the third part of this article.)

Graphics Can Lead to More Thorough Data Analyses

The best graphics show data sets as a whole while also displaying the larger trends, which leads to more thorough data analyses (Cleveland, 1985). Tufte (2001) claimed that graphics can be designed to have three layers—one showing the overall structure of the data (a summary), another the close-up details (the raw data), and the last an implicit answer to the question underlying the graphic (inference), such as 'which group performed better?' or 'is there a relationship between these variables?'

Tufte (1990) gave the example of the Vietnam Veterans' Memorial in Washington, DC, as an example of a design that achieves visual and emotional power by using several layers of visual information. From far away, the wall, which lists the names of the 58,000 American soldiers killed in the Vietnam War, conveys a strong visual impact of the huge cumulative toll in lives (the summary). Coming closer to the wall, the individual names of the dead can be made out and examined (the raw data). Tufte (1990) stated: "Thus, the names on stone triple-function: to memorialize each person who died, to make a mark adding up to the total, and to indicate sequence and approximate date of death" (p. 44).

Likewise, researchers can also reveal the power of their data at both a macro and a micro level by using data accountable graphics. Tufte (1990) argued that these types of displays give control

of the information to the viewers, thus adding to the credibility of the argument the author is making. In other words, the graphics containing the details will allow the reader to ascertain just how far the summary statistics accurately represent the data. Data accountable graphics let viewers make their own independent evaluation of the data. A graphic that shows both the details and the overall structure of the data helps the reader to "retain the information in the data" (Deming, quoted in Cleveland, 1985, p. 9) much better than a summary statistic.

Through it all, the best graphics will lay out data in a way that lets readers compare some variable to another. Tufte (1997) summarized matters like this: "The deep, fundamental question in statistical analysis is *Compared with what?*" (p. 30).

Graphics Provide a Compact Way to Make All the Data Available to the Reader

A typical journal article contains many details about the methodology, participants, and instruments used in a study. It is, then, anomalous that the data upon which important statistical conclusions are founded are not also available. Graphics provide a way to present all the data in a compact format, doing so in a way that should help readers to see the pattern in the data for themselves. For example, a scatterplot provides a way to ascertain the original data points but also gives the big picture of the trends and possible distributional anomalies in the data. Larson–Hall and Plonsky (2015) have recommended that data rich graphics should be included in almost all studies and that raw data should be shared whenever possible.[3]

Weissgerber et al. (2015) insist that data accountable graphics are even more desirable when sample sizes are small (around $n = 10$) as summary statistics are not very meaningful in such cases. Their graph, reprinted in this article as Figure 1, helps make this case.

Graphics Can Be Used to Provide a Quick Idea of the Topic of the Research Article

As already mentioned, journal readers may find graphics the least intimidating part of a research article and be attracted to look at this material first as a way to scan whether the topic of the article is suitable for further reading. Another intriguing trend is graphical abstracts. These are meant to help readers understand the topic of the research article in just a glance and thus "encourage browsing, promote interdisciplinary scholarship, and help readers quickly identify which articles are most relevant to their research interests" (Alves, 2013). Figure 2 shows a graphical abstract cited as exemplary on the Elsevier website ("Graphical Abstracts," 2015) that makes that point.

Although the graphical abstract in Figure 2 does not pertain to the L2 research field, making it hard to evaluate it independently, it suggests additional possibilities offered by visualization techniques that might be worth exploring. As a follow-up to that idea, I have therefore created a sample second language research graphical abstract concerning the relationship between input and output in bilingual children in the research by Pearson and Fernandez (1994). The graphic (see Figure 3) quickly shows that in a bilingual child's development, the amount of output in each language is highly affected by the language use that the child encounters in each language.

To sum up, I have argued that data accountable graphics can enhance research reports because they are attractive to readers, may reveal patterns in the data that would not have been seen without visuals, and make good use of the brain's prodigious capacity for visual understanding, which is often unconscious and automatic. These attributes of graphics mean that we may understand graphics more quickly, holistically, and easily than numbers, leading to more thorough and potentially more informative data analysis. The use of data accountable graphics also makes the reader an active participant in the scientific process by encouraging critical thinking about the author's analysis (Weissgerber et al., 2015).

A HISTORICAL VIEW OF THE USE OF GRAPHICS IN L2 RESEARCH

Within L2 research, there has clearly been a progression in the frequency and sophistication of quantitative approaches to data. Loewen and Gass (2009) outlined this trend for statistical tests by noting a substantial increase in the number of articles containing inferential statistics beginning in the 1970s. My interest was in determining whether a similar increase in sophistication might also hold for the use of graphics. Although it is true that there is not only one 'correct' way to present any given data visually (Tukey, 1977), it is also true that not all graphics are created equal (see Larson–Hall & Herrington, 2009). Some graphics provide much more information and are much more valuable in terms of the space they occupy on the page than others. Tufte (2001) praised the scatterplot as one of the "greatest of all graphical designs" (p. 47), Cleveland (1985)

FIGURE 2

Reprinted From Journal of Controlled Release, 154(3), Tomáš Etrych, T., Kovář, L., Strohalm, J., Chytil, P., Říhová, B., & Ulbrich, K., Biodegradable Star HPMA Polymer–Drug Conjugates: Biodegradability, Distribution and Anti-Tumor Efficacy, p. 241–248, Copyright (2011), With Permission From Elsevier. [Color figure can be viewed at wileyonlinelibrary.com]
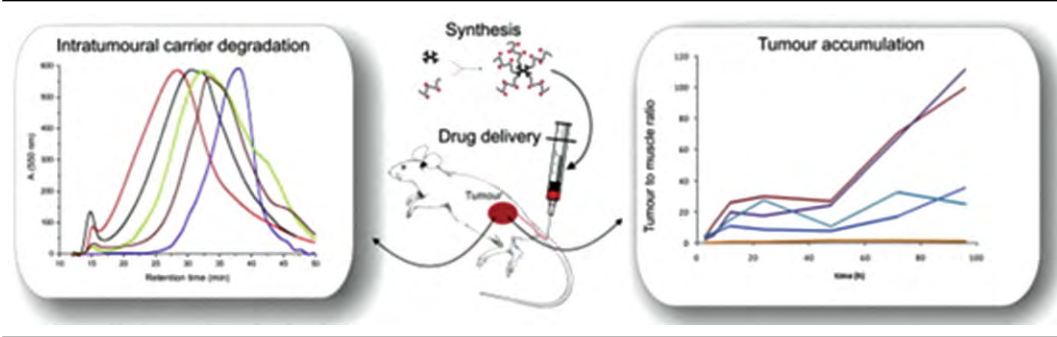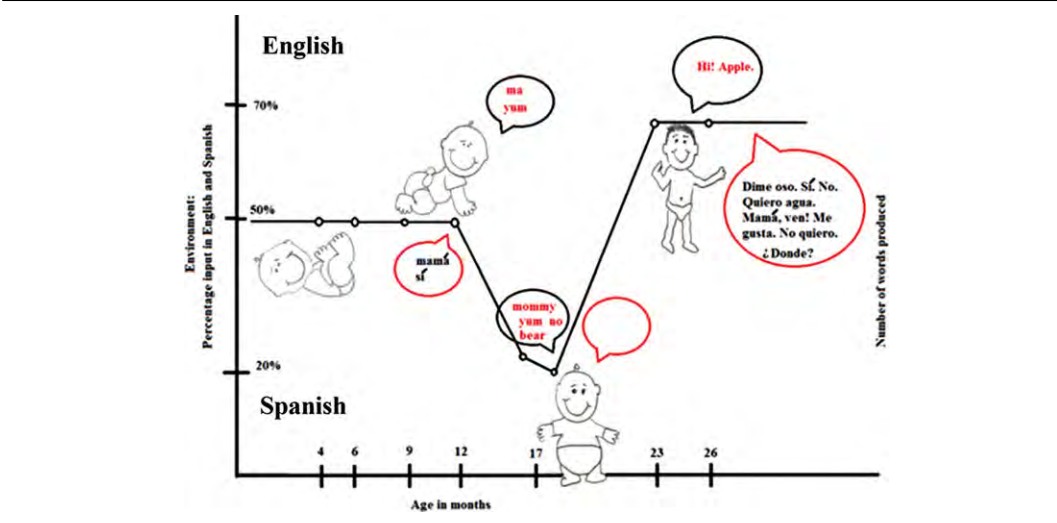

FIGURE 3

A Possible Graphical Abstract for Pearson & Fernandez (1994) [Color figure can be viewed at wileyonlinelibrary.com]

invented the Loess line that can be overlaid on scatterplot data, and Tukey (1977) created the boxplot and variations of it. With that as background, I was looking to see whether graphics had evolved from simpler to more complex, informative, and data accountable graphics.

I surveyed three L2 research journals from their beginning date of publication through 2011 or 2012 and classified the type and number of graphics included in empirical articles (total number of empirical articles surveyed = 1835). The journals were *The Modern Language Journal* (*MLJ*), which began publication in 1916; *Language Learning* (*LL*), which began in 1948; and *Studies in Second Language Acquisition* (*SSLA*), which began in 1978.

My research questions were the following:

RQ1. Did the percentage of empirical articles that contained graphics increase over time?

RQ2. What percentage of graphics were line graphs, bar plots, scatterplots, path diagrams, or other types, and did this change over time? In this count I was not interested in the total number of graphics found in each article but rather the types of graphics.

RQ3. How many articles included either a data accountable graphic, from which the full data set could be recovered, or the actual data?

The only methodology employed in this study was to physically (or virtually, online) flip through

every page of every issue of the three journals for their entire history and note which articles contained empirical data, what graphics were present at least once in any given article, and whether data accountable information was available. I defined progress in the use of graphics as a higher percentage of empirical articles containing graphics and a greater use of data accountable and data rich graphics.

General Results

For all of the research questions, it was surprising to discover that there has been as much regression as there has been progress. Before the 1970s, few empirical studies were published at all, but many of those that were contained graphics, usually hand-drawn, or provided the raw numbers used to calculate the inferential statistics. This may have been because authors did not expect their readers to understand the meaning of inferential statistics very well. Lemper (1925), for example, noted that the article would offer data from "objective" and "subjective" tests by raters in a visual way by presenting individual scores rather than by giving a correlation coefficient, because it would "be intelligible for a larger number of readers" (p. 176). Full individual data were listed for the 28 participants.

RQ1   Did the percentage of empirical articles that contained graphics increase over time?

Figure 4 visually summarizes the answer to RQ1; it shows the percentage of empirical studies per year which contained graphics in the three journals surveyed. Figure 4 uses sparklines, a small multiple graphic proposed by Tufte (1990, 2006). These sparklines are similar to bar plots for count data, except that they are very small, which makes the actual data points unrecoverable (however, the Excel file that contains the actual counts has been made available online at the Inter-University Consortium for Political and Social Research (ICPSR) database; see Larson–Hall, n.d.).[4] The point of the sparkline is to put together large amounts of data in a very small space so that trends can be discerned (Bissantz sparklines were used here as the resolution of sparklines available in Excel are not as sharp). Numbers before and after the sparklines under the dates note the actual percentage of studies containing graphics at the beginning and end of the period, with the gray (highlighted blue in the color version) bar indicating the high in the data. For *MLJ* and *LL*, this gray bar is the first tall bar encountered moving from the left and is the maximum 100%. For

SSLA, the high of 90% is the gray bar in the middle of a surge of graphics in the mid-1990s. For all of the journals the low is 0%. The highest bars, where a larger percentage of the articles contained graphics, are found in the earlier years of the journals. For *MLJ* and *LL*, the tallest bars represent 100% of the empirical studies; but these are only found before the 1970s.

From the 1980s to the present there has been some increase in the number of graphics published, although fluctuations can be large. Still, graphics seem to be more routinely included in articles within that period. *MLJ* and *LL* have included a similar number of graphics: *MLJ* has fluctuated from a low of 0% to a high of 56% of empirical articles containing graphics (average 31%), while *LL* has fluctuated from a low of 7% to a high of 71% (average 39%). By contrast, *SSLA* has shown an increase over time, from a high of only 33% of graphics in articles published between 1978–1987, to highs of 90% from 1988–1997 and 85% in the years between 1998–2011, with an average of 53% (from 1988–2011).

RQ2   What types of graphics were used historically?

Table 1 summarizes the types of graphics used over the entire history of each journal. This is a count of articles that included at least one instance of the given graphic; if an article contained a boxplot, it was counted as one instance of a boxplot, even though the article may have included more than one boxplot.

Table 1 shows clearly that bar plots and line graphs were the graphics used most often over the entire history of all three journals, found in at least 70% of the articles containing graphics for each journal. The use of other graphics is truly marginal.

However, there has been a notable change in one area. Before the year 1970, the total number of empirical articles was quite small. Although, in terms of percentages in a single issue, a larger number of these articles included graphics than today (see Figure 4), the total number of graphics nevertheless was very small. From 1916–1969, *MLJ* published 222 empirical articles, and 24 contained graphics of any kind. From 1948–1969, *LL* published 18 empirical articles, and 9 had graphics (*SSLA* is not mentioned because it did not begin publication until 1978). In these 33 articles, there were 36 examples of graphical types: line graphs (16), bar plots (9), boxplots (4), scatterplots (2), frequency polygons (2), histograms (2), and regression lines (1). Although line graphs and bar plots made up the majority of the

FIGURE 4

Counts of the Number of Empirical Articles Which Contained Graphics in Three Second Language Research Journals [Color figure can be viewed at wileyonlinelibrary.com]

TABLE 1

Types of Graphics Used in the History of Three Journals, Calculated by Number of Articles Including at Least One Token

	MLJ (1916–2011)	*LL* (1948–2012)	*SSLA* (1978–2011)
Bar plot (% of total)	83 (39%)	110 (34%)	107 (46%)
Line graph (% of total)	79 (37%)	140 (43%)	95 (41%)
Scatterplot (% of total)	13 (6%)	22 (7%)	19 (8%)
Path diagram or similar map of statistical relationships (% of total)	9 (4%)	22 (7%)	2 (1%)
Other type of graphic (% of total)	27 (13%)	33 (10%)	8 (3%)
TOTAL:	211	327	231

graphics, all of the other categories represent a fair number of data rich and data accountable graphics (11 out of 36, or 31%).

During the 1970s, the predominant graphic in *MLJ* and *LL* was the line graph, with 25 of the 34 examples (74%) from this time period being line graphs (only 3 contained bar plots and the other 6 were neither). From about 1987, bar plots also became common, with these two graphics (line graphs and bar plots) comprising the majority of included graphics to the present day.

The results found in Hudson (2015) show that the pattern of the field using line graphs and bar plots as the predominant graphics still holds today. Hudson examined the number of visuals used over one recent year of time in five L2 journals and found 136 empirical research articles with 207 graphics included. Of these, 50 were instances that did not display actual data as they were diagrams using text only, pictures (assuming this means photographs), or spectrograms. Of the remaining 157 graphics documented by Hudson, 103 (66%) were line graphs or bar charts.

Surveys of other scientific fields show that this pattern is not at all unusual. In the field of Psychology, the majority of graphs are bar charts (Sandor & Lane, 2007). A survey of 703 articles in top Physiology journals found that most graphics showed continuous data in bar plots and line graphs: 86% of articles included at least one bar plot, while line graphs and point and error bar plots appeared in 61% of the articles (Weissgerber et al., 2015). Only 13% had at least one scatterplot, 5% had one or more boxplots, and 8% had at least one histogram. In the field of medicine, 63% of the 56 articles surveyed in 6 issues of the *Journal of American Medicine* used bar plots or line graphs (Cooper, Schriger, & Close, 2002).

RQ3 How many articles were data accountable?

Only a very small percentage (7%) of the total number of articles published in the L2 research field have provided either a data accountable graphic, from which the original data points could be extracted, or an actual data set underlying the statistical conclusions drawn in the article. By contrast, empirical articles published

before 1970 were quite exemplary in this regard, with 55% being data accountable. For *MLJ* before 1970, 24 empirical articles were published and 14 of those either contained the individual raw data printed with the article (11), or a data accountable graphic. In *LL* before 1970, 18 empirical articles were published and 9 of those were data accountable (8 with raw data, 1 with a scatterplot).

While the 1970s saw a great increase in the number of published articles which were empirical, data accountability practically disappeared. For example, in *LL* from 1970–1979, 189 articles were published, of which 103 were empirical, but no articles published raw data, and no graphics were data accountable. In *MLJ* between 1970–1979, 69 empirical articles were published but no raw data was provided, and there was only one data recoverable graphic in this period (a scatterplot).

For more recent times in the history of the journals the percentage of data accountable graphics has continued to be very small, though slightly better than the 1970s period. From 1981 to 2012, *MLJ* published a total of 452 empirical articles and only 4 provided raw data, 12 provided a data accountable graphic (all scatterplots), and there were only 6 articles where data rich graphics (boxplots and histograms) were provided. *LL* showed a similar pattern with 557 empirical articles from 1980–2012, but of these, only 11 had raw data, 21 contained scatterplots, 5 more had data accountable graphics of other kinds (a frequency plot, a recurrence plot, and small multiples), and 8 more instances were of data rich graphics that showed underlying distributions (boxplots or density curves). *SSLA* fared the best in providing data accountability over its history from 1978–2011, with 380 empirical articles published, 16 of which contained raw data; even so, this amounts to only 4% of the articles, a tiny fraction and nowhere near comparable to the high percentage found before in the journal articles published before the 1970s. There were 20 instances of data accountable graphics (scatterplots and 1 histogram) and 6 more instances of data rich graphics (boxplots and 1 small multiple), all published only since 2004.

Conclusion

Historically, what I have found is that before the 1970s, when the field was young and not many precedents (or software) existed for creating graphics, authors more often used data accountable types of graphics and/or provided entire datasets in their articles. A striking change in the L2 research field came during the 1970s, when more empirical articles began to be published and more graphics were produced, but data accountability in the field quickly diminished. The field settled on the bar plot and the line graph as consensus graphics.

MOVING TOWARD BETTER GRAPHICS

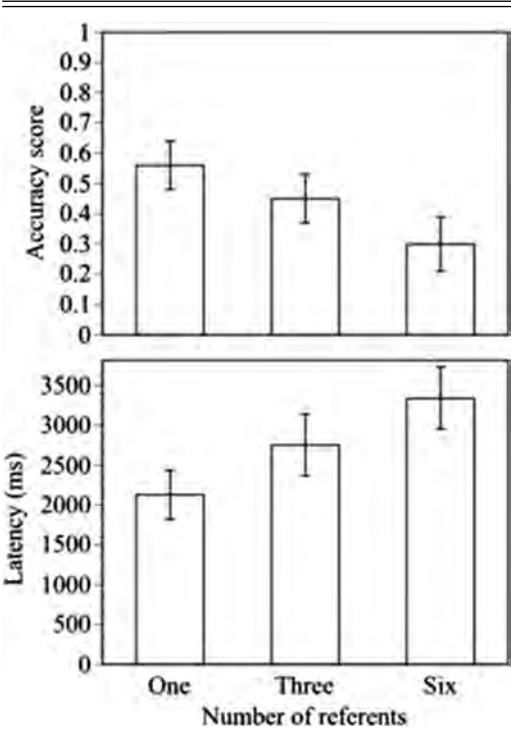### *What's Wrong With Bar Plots and Line Graphs?*

The bar plot is a venerable graphic, first seen in Playfair's designs from 1765 or 1785 (Tufte, 2001). Friendly (2006) documented the appearance of line graphs in the early part of the 18th century. Thus, both forms have been around a long time and have been able to gain a strong following. But this does not mean that they are useful tools. Unwin (2008) noted that "[w]ith the advent of computers [in the 1970s], graphics went into a relative decline" (p. 60). While as graphics they became much easier to draw they were "simple and rather ugly" (Unwin, 2008, p. 60).

More seriously, in plotting continuous data, both bar plots and line graphs are severely lacking in the information they present. To begin with, "[t]he [bar plot] wastes space; you could show at least 100 numbers in the space that now shows 1 number" (Tufte, 2014), or, worded differently, "[i]t is certainly profligate to use an entire bar when all of the information about the mean is contained in the location of the top line; the rest is chartjunk" (Wainer, 1996, p. 105). Furthermore, for interval-type data, the bar plot and line graph do not show much data and thus do not take advantage of the human ability to visually make sense of complex data.

By now, many researchers add error bars to their bar plots and line graphs, resulting in three points of information for one group rather than just the one point of the mean score. But this actually complicates matters inasmuch as those little bars could be one of three different types: standard deviation bars, standard error bars, or 95% confidence intervals. Cumming (2012) notes that authors must be careful to label the bars in their figures or readers will not know which kinds of error bars are being used and may misinterpret them. Figure 5 is a bar plot used to plot continuous data by Sommers & Barcroft (2013), which appropriately labels the bars in the figure description as "Accuracy (Top) and Latency (Bottom) for the Picture-to-L2 Recall Task; Error Bars Represent Standard Deviations of the Mean." Without that description the reader could not determine which kind of bars were being used.

FIGURE 5

Sommers & Barcroft (2013) Bar Plot With Error Bar Labeled, With Permission From Wiley

A study by Belia et al. (2005) found that most of their survey respondents—scholars with published articles in the fields of psychology, behavioral neuroscience, and medicine—could not correctly interpret the inferential meaning of overlap or non-overlap of standard error bars or CIs. This implies that error bars are not a good way to establish visually whether groups are statistically different or not. A further problem is that error bars cannot be used at all to make statistical inferences between pretest–posttest measures (repeated measures) as they fail to take into account the correlation between the two measures.

In a graph like Figure 6 from Saito (2013), the error bars are appropriately labeled as CIs (the figure's original title reads "95% confidence intervals and mean values of the learners' pre-post test scores.") and Saito did not attempt to draw any statistical inferences about whether there were differences between the pretest and the posttest. However, the fact that CIs were added to the figure of the development data may invite readers to think that such inferences can be drawn.

What kind of graphics would be able to show statistical differences between groups? Because boxplots also do not contain information allowing any statistical comparison between groups, they would not solve the problem. CIs plotted separately or imposed over a boxplot or univariate scatterplot are possible; in that case it would be prudent for authors to state the inferential conclusions that one can draw from them, perhaps in the figure label. An alternative is to put data into an Excel program called ESCI created by Cumming (detailed in Cumming, 2012) that can plot CIs on the same chart as the data. An example for data from Poehner & Lantolf (2013) on how mediation, given as hints for understanding, helped improve scores on a Chinese listening–comprehension test, is shown in Figure 7 (from Larson–Hall, 2015a, p. 139). The left side of the graph shows the descriptive data while the right side shows difference data and a CI of the difference data, allowing for the conclusion that the two groups of scores are statistically different. More specifically, the lines show the actual paired points of the participants. The two black lines with circles in the middle show the CIs of the two sets of data ("Actual" and "Mediated"), and the triangles represent the gain scores (the difference from "Actual" to "Mediated" scores) while the black line with the triangle in the middle shows the CI for the data differences, with the scale of differences listed on the right of the figure. Figure 7 is thus data accountable but also shows inferential statistics. Due to data differences, CI does not run

However, Sandor and Lane (2007) point out that, even if bars are labeled, they simply say something about the variability within the group but do not easily let the viewer make inferences about *differences* between groups unless the reader knows how to understand the overlap or non-overlap of the bars. For example, readers may assume that, if the error bars overlap, there is no statistical difference between groups; but this is false. First of all, standard deviation bars are simply descriptive and cannot be used to make inferences at all (Cumming, Fidler, & Vaux, 2007). Confidence intervals (CIs) and standard error bars *can* be used to infer statistical differences between groups. Both CIs and standard error bars show a region where one can expect to find the true mean. In interpreting inferential statistics, for CIs, bars may overlap up to about 25% (or even more if sample sizes are smaller than 10) of their average length and still retain statistical differences, while for standard error bars there must be no overlap and a gap between the ends of the error bars that is equal to at least one half of the average length of the error bar, if the sample size is at least 10 or more (Cumming et al., 2007).

FIGURE 6

Graph From Saito (2013) With Repeated Measures, With Permission From Wiley
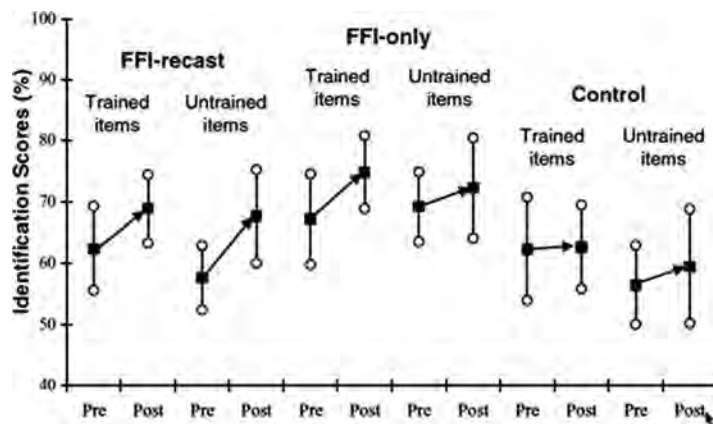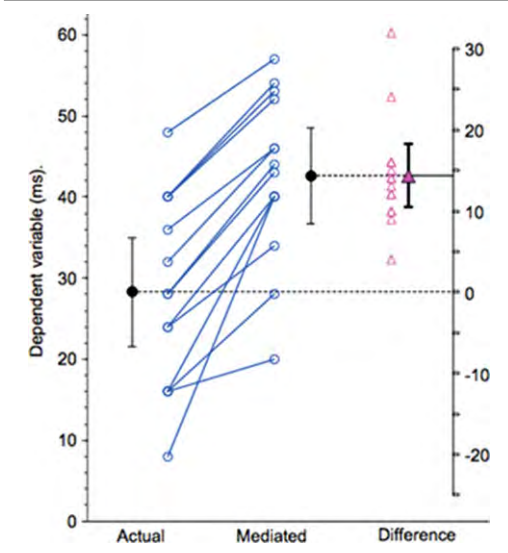

FIGURE 7

Graph Showing Confidence Intervals Plotted on a Univariate Scatterplot With Two Paired Variables, Allowing Statistical Inference, With Permission From Taylor & Francis [Color figure can be viewed at wileyonlinelibrary.com]

through zero, there was a statistically significant difference between the groups, which means that the mediation helped improve scores. Additionally, the CI runs from 10–18, meaning that the amount of help the mediation provided can be estimated to be somewhere between 10 to 18 points of a boost to the test score.

Another problem with bar plots and line graphs for continuous data is that they are not data accountable, nor are they even data rich. They do not show information about the shape of the distribution of the data, which makes understanding the true nature of the data impossible and can significantly change the conclusions drawn from the data using only summary information (see Weissgerber et al., 2015, in reference to Figure 1, and the attendant discussion). We cannot see the performance of individuals in the graphic, only the overall trends. Tukey (1977) pointed out that "it very often pays to plot residuals" (p. 125). By this he meant that we look at the overall trends and compare how the individual points relate to it. This, however, is not possible with the bar plot or line graph.

When Are Bar Plots and Line Graphs Appropriate?

Bar plots do have a place in research reports. They are useful for comparing counts, as opposed to mean scores. Figure 8 is a bar plot from Gullberg (2006), which shows counts of how references were grammatically encoded in L1 and L2. Originally entitled "Mean proportion of instances of maintained reference encoded as NP Lex, NP Pron, or NP Ø in L1 and L2," this figure actually shows the proportion of the references that were encoded by these different grammatical means, and this is essentially a number of counts, just standardized. When bar plots are used to count things, no error bars are possible because there is no variation, only the count of things. With data where the exact count is not so important but comparing counts over time or a large number of categories is, the sparkline might be a useful graphic (as seen in Figure 4; see also Hudson, 2015, for an example). Figure 9, reworked from Clark & Clark (1966), shows how counts can be

FIGURE 8
Appropriate Use of Bar Plot, Gullberg (2006) Focused on Counting Number of Instances, With Permission From Wiley
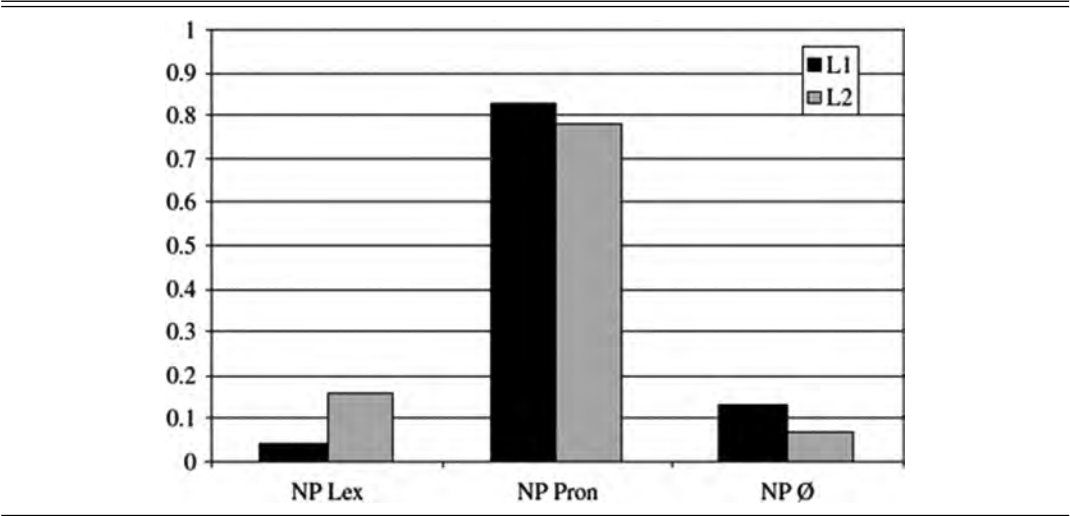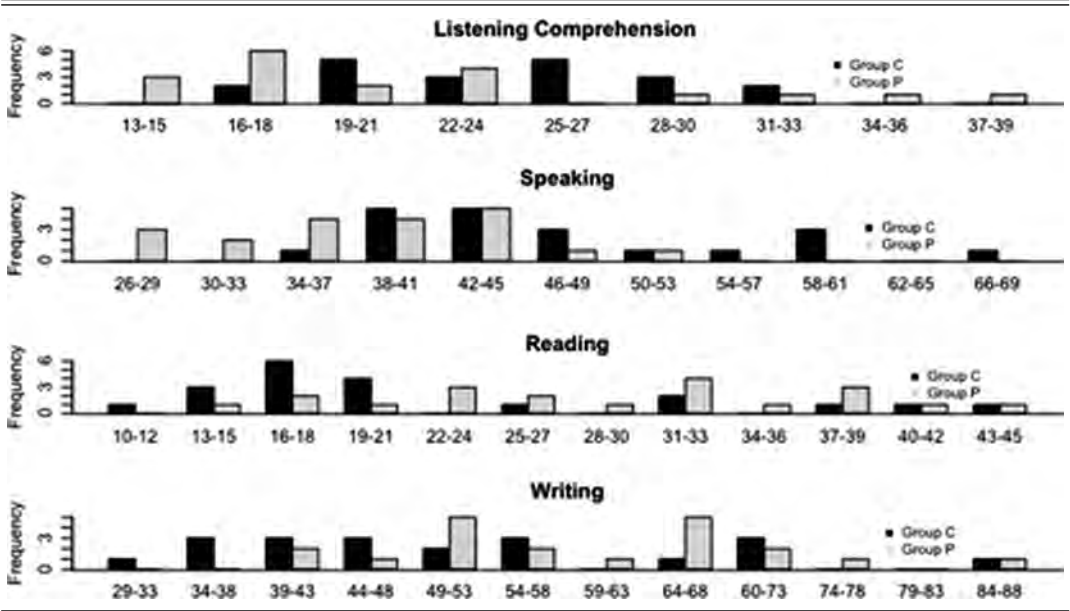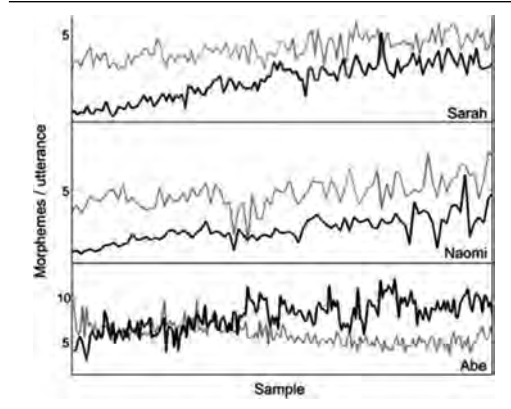

FIGURE 9
Figure by Clark & Clark (1966), Exemplifying Condensed Bar Plots, With Permission From Wiley

compressed into a smaller space when there is a lot of data. Kampstra (2008) stated that comparing multiple histograms may be difficult because of space concerns, but Figure 9 overcomes this problem by reducing the size of the histogram, while still being larger than the sparkline. Notice that both Figure 8 and 9 are data accountable figures because they provide all of the data points in the data set.

Re-created for better resolution from Clark & Clark (1966), Figure 9 is an example of condensed bar plots and is similar to histograms. Its original title reads "Distribution of test scores" (see the Appendix for R Code).

When the data are plentiful, line graphs can be a good way to display data while providing data accountability. Consider Figure 10, which was printed in Dale & Spivey (2006, p. 410) and

FIGURE 10
Dale & Spivey's (2006) Graphic Showing Average Morphemes per Sentence for Three Children Acquiring a Language Over Time, With Permission From Wiley

shows the longitudinal morphemic development of three children. Its original title provided this specific information:

> *Figure* 6. Abe, Sarah, and Naomi (black lines) exhibit different levels of development in terms of the number of morphemes per sentence in the transcripts (gray lines present these data for caregivers). Values were calculated by dividing the total number of syntactic elements used by the number of sentences in a transcript.

Here, the graphic is certainly data accountable; the three children's data are reprinted next to each other so their development can be compared, and the caretakers' lines are also given for comparison.

Graphics for Presentation Versus Graphics for Data Exploration

Most of this article addresses graphics for presentation or the graphics that are printed in research articles. However, graphics are often quite useful to a researcher well before any publication as a means of exploring data. They retain all of the advantages mentioned in the first part of the article—they help the researcher find patterns that may not be noticed in summaries, including outliers or deviant data entries. They also lead to more thorough data analysis as the researcher is able to explore the data, by means of data accountable graphics, at both the macro and the micro level, looking at summary trends as well as individual performance. Unwin, Chen, and Hardle (2008) asserted that the graphics available in computer programs are much more useful for

obtaining quick and plentiful exploratory graphics than the more careful presentation graphics, although no statistical software, including SPSS, SAS or R, provides "effective interactive tools for exploratory graphics" (p. 6). In other words, sometimes it is not easy to draw the graphic that you want with the default choices in software programs, much less uphold rules of good graphical presentation (see Hudson, 2015; Tufte, 2001; Unwin, 2008).

That said, here are some recommendations for exploring continuous data sets, depending on whether they are testing differences in groups or looking for relationships among variables. In both cases, a good place to start is Atsushi Mizumoto's langtest.jp Web site, which provides a number of exploratory graphics along with statistical conclusions.

Group Differences With Two Groups. Mizumoto's Web site offers overlaid histograms, allowing comparison of the distribution of the two groups. Outliers, skewed distributions, and multimodal distributions (distributions with more than one high point) can be easily spotted with histograms. Boxplots overlaid with individual data points (beeswarm plots) help visualize the forest *and* the trees and should help the researcher get a sense of whether there are real differences between groups. Outliers are also clearly labeled (see Figure 12 for an overlaid histogram and beeswarm plot). The pirate plot is also an excellent plot for visualizing distribution and summary statistics at the same time (see Figure 14 for an example).

Group Differences With Three or More Groups. Unfortunately, the graphics on Mizumoto's website are not as useful for group comparisons with more than one group or more than one variable. In the ANOVA tab there is one plot of mean scores with CIs. If you use this, it is critical to keep in mind the rules about interpreting statistical group differences mentioned earlier. Parallel coordinate plots are good for repeated measures data where participants' responses are measured more than once.[5]

Relationships Between Two Variables. Mizumoto's langtest.jp website provides a beeswarm plot, histograms with overlaid density curves, and robust scatterplots (where an ellipsis encircles the 'good' part of the data). The scatterplot is especially important as it will visualize what kind of relationship the variables have, but also show any bivariate outliers (outliers that can only been seen at the intersection of the two variables). If the robust scatterplot only includes a small sample of

the data it could indicate a nonlinear relationship in the data, which may suggest splitting the data by another variable. Scatterplots overlaid with Loess lines are useful for easily showing different kinds of distributions between the two variables, such as strong correlation, no correlation, and nonlinear relationships such as curves in the association line.[6]

Relationships Between More Than Two Variables. Scatterplot matrices are useful for looking at more than one relationship at a time, and langtest.jp provides an easy way to generate such a plot, which has histograms overlaid with density curves for individual variables on the diagonal.[7]

Non-continuous Data Sets. In case you are using data that is all categorical, there are a number of really interesting new visualization techniques for this kind of data available in R; one of them, a mosaic plot, can be found on langtest.jp in the chi-square choice, and several more are described in the online chapter by Larson–Hall (2015a) about Chi-square, available at http://routledgetextbooks.com/textbooks/_author/larson-hall/ under "Supplemental Material."

### What Can Be Used Instead of a Bar Plot or Line Graph?

In creating presentation graphics for data that is not count data but instead has categorical groups, such as an experimental versus a control group and resulting scores, there are a number of alternative graphics that are data accountable or data rich and, in addition, simply more interesting to look at than bar plots.

Graphics for Groups. One of the main alternatives to the bar plot or line graph is the boxplot, first laid out in Tukey (1977). The boxplot shows five pieces of data: the median, the ends of the first and third quartiles of data which are contained in the box, and the upper and lower whiskers which extend out to the minimum and maximum values in the data unless they are deemed to be outliers, which are shown with small dots. The boxplot is a data rich graphic. However, overlaying the individual data on top of the boxplot is an addition that also makes the boxplot data accountable: It becomes more similar to the scatterplot in that all of the data can be seen, while also providing a way to see the bigger picture of the data distribution. Figure 11 shows a boxplot with overlaid individual data (Larson–Hall & Connell, 2005). Figure 11 shows that the B and

FIGURE 11
Larson–Hall & Connell (2005), Showing a Boxplot With Overlaid Dots (a Beeswarm Plot), Illustrating a Data Accountable Graphic
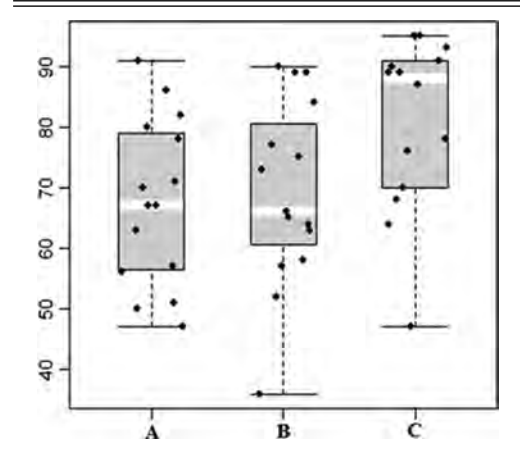

C groups have very large whiskers on the boxes. But because the dots are overlaid, the figure also shows that there is only *one* participant that has caused the whisker of both of these groups to go so long. The dots below the 50% percentile box are clustered in close, but for both groups there is one dot quite far away. The point is that a data accountable graphic enables readers themselves to see what is happening in the data, making them active participants in evaluating whether the data are skewed or symmetrically distributed, how strong the trends are, and so on.

This graph could be further enhanced with the addition of a mean score line superimposed over the boxes and 95% CIs also inserted. This type of graphic, a so-called beeswarm plot, can be accessed online at langtest.jp by following the links for "Comparing two independent samples" if the data are from separate groups and "Comparing paired samples" if the data are from the same participants. To create Figure 12, I used data from Abrahamsson and Hyltenstam's (2009) study on nativelike pronunciation for Spanish L1–Swedish L2 speakers divided into two groups (with age of onset from 0–11 years in Group 1 and 12 and older in Group 2). Both the overlaid histogram on the left side of the graphic and the beeswarm plot on the right were automatically generated (I changed labels with Photoshop). In this study the score that each individual received was based on whether each of 10 judges rated the person as a native speaker or not (a simple binary choice for each judge, but totaled together for the perceived nativelikeness score). Beyond the points that are superimposed on the boxplot, the mean

FIGURE 12

Abrahamsson & Hyltenstam's (2009) Data Using an Overlaid Histogram (Left Side) and Beeswarm Graphic (Right Side) With Mizumoto's langtest.jp Web page Defaults [Color figure can be viewed at wileyonlinelibrary.com]
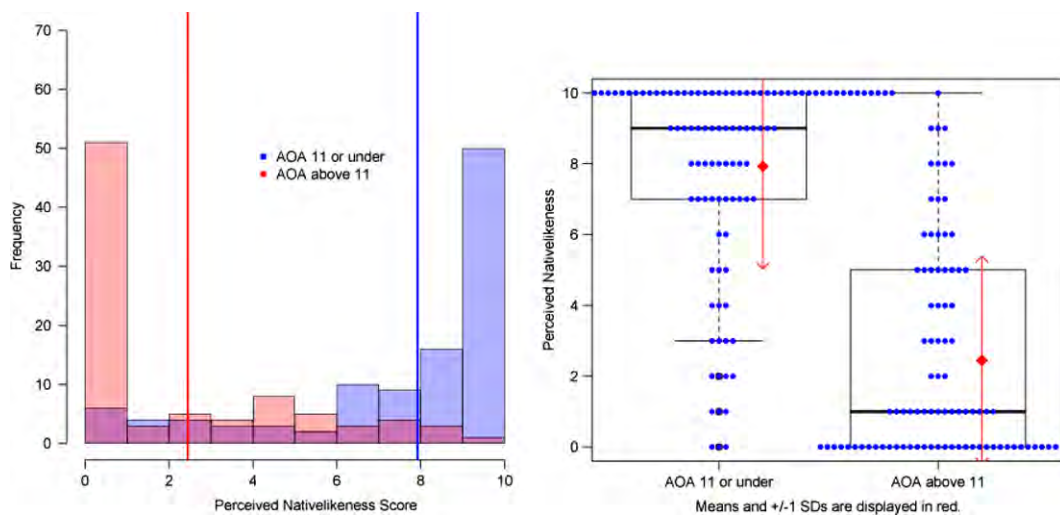
and CI are also shown with the gray diamond and line (red in the online version) through the diamond. Although, as noted by Abrahamsson & Hyltenstam, age 11 is only a theoretically appropriate age for dividing this data, the beeswarm plot shows clearly that the great majority of cases are found within the boxplot portion of the data and also shows why the statistical outcome of the independent samples *t*-test is that the groups are different from each other. In the graphic readers can see the group trends but can also pick out information about exceptional individuals. For example, for those who started learning Swedish above age 11, the number of individuals who passed as native speakers in the judgment of 7, 8, 9, or 10 native judges was quite small but not beyond the upper whisker, so they are not considered outliers. On the other hand, for individuals who began learning Swedish at age 11 or lower, those who did not pass as native speakers in the eyes of at least three judges *are* considered outliers because they are below the lower whisker of the boxplot.

Unfortunately, Mizumoto's Web site does not have any way to plot more than two groups at this time, but using Mizumoto's R code as a help, I created Figure 13, which shows the beeswarm plot data for all three groups in Abrahamsson & Hyltenstam (2009), and additionally plots a mean dot and standard deviation error bars on top of that figure (for R code see the Appendix).

Another way to achieve the same effect in an easier fashion, although without the mean dot and error bars, is to use the website http://data.vanderbilt.edu/~graywh/dotplot/. Here, data for any number of groups can be pasted in to generate the beeswarm graphic.

A recently invented graphic, shown in Figure 14, is the pirate plot (Phillips, 2016). Instead of the boxplot structuring the data, the pirate plot uses a beanplot, which shows the density curve of the data (like a histogram, just smoothed over) symmetrically around the vertical axis. Kampstra (2008) pointed out that boxplots are not intuitively understandable, and that they may mask unusual distributions like those with two modes. In addition, trying to show distributions with boxplots or stem-and-leaf plots is difficult because of the space they take up. Note that the overlaid boxplot generated by Mizumoto's site in Figure 12 helps alleviate the space issue, but the beanplot, which is more visually intuitive, also takes only the same amount of space as a boxplot. The beanplot (not shown in this article) also displays all of the individual data points as lines representing peas within the pod of the bean (plot). The pirate plot builds upon the good points of the beanplot, such as the use of raw data and intuitive description of the data distribution by means of the smoothed density curves and mean points, and adds inferential data in the form of a CI band (the R code for

FIGURE 13

Abrahamsson & Hyltenstam's (2009) Data in a Beeswarm Graphic With 3 Groups, Using the R Statistical Program [Color figure can be viewed at wileyonlinelibrary.com]

Means and +/-1 SDs are displayed in red.
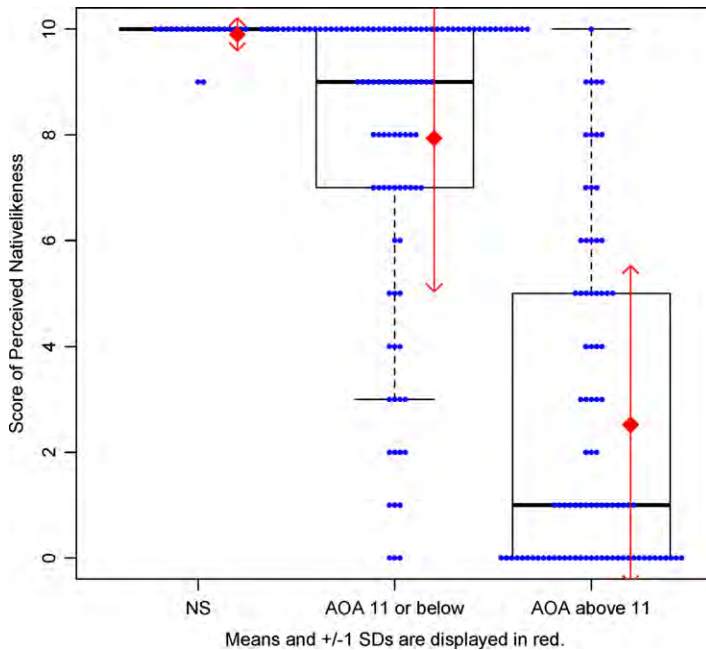
this plot can be found in the Appendix). Phillips (2016) preferred to add a Bayesian inferential band called the 95% Highest Density Interval but also provided a 95% CI, which is what can be seen in Figure 14. Interpreting the 95% CIs in Figure 14 is easy, as they are spread quite far apart and clearly show that the groups are statistically different. They also show that the CI is much smaller for the native speakers than for the other two groups.

The original graphic in Abrahamsson & Hyltenstam (2009) is also a data accountable graphic, and is reprinted in Figure 15. Originally entitled "Scatter plot of PN [Perceived Nativelikeness] scores versus AO [age of onset] for all 195 participants and the 20 native controls (AO 0 Years)," it is similar to a stem and leaf plot in that numbers are used to show the distribution.[8] What is attractive about this plot is that, although the authors make a theoretical division into younger and older learners between ages 11 and 12, the full display of their data enables readers who might want to look for other places where a division might make sense, to do so.

The different options that I have provided for Abrahamsson & Hyltenstam's (2009) study illustrate well that many kinds of graphics can be

used to examine and present the same dataset. In fact, Cleveland (1985) claims that looking at a variety of visual configurations of data is beneficial as researchers consider their results. Such an assessment is all the more important since it will not be possible to publish all the graphics; in other words, a decision must be made which graphic(s) best fits the information that is to be conveyed and is likely to make the most sense to readers.

Graphics for Repeated Measures. Boxplots, beeswarm plots, and pirate plots can also be used with data that are related in some way, such as data from the same groups at different time periods, or data from the same groups using tests broken into related parts (such as a test of past tense and another one of present tense or different phonological categories). However, another useful graphic is the parallel coordinate plot. Mizumoto's langtest.jp Web site will automatically create a parallel coordinate plot in response to the entry "Comparing paired samples," although this will only work for two sets of data at a time. Using Mizumoto's Web site I entered data for two groups I was examining to see how use of a bilingual computerized dictionary program would affect the number of corrections students made to

FIGURE 14

Pirate Plot of Abrahamsson & Hyltenstam's (2009) Data on Number of Judges Who Considered Swedish L2 Speaker a Native Speaker [Color figure can be viewed at wileyonlinelibrary.com]
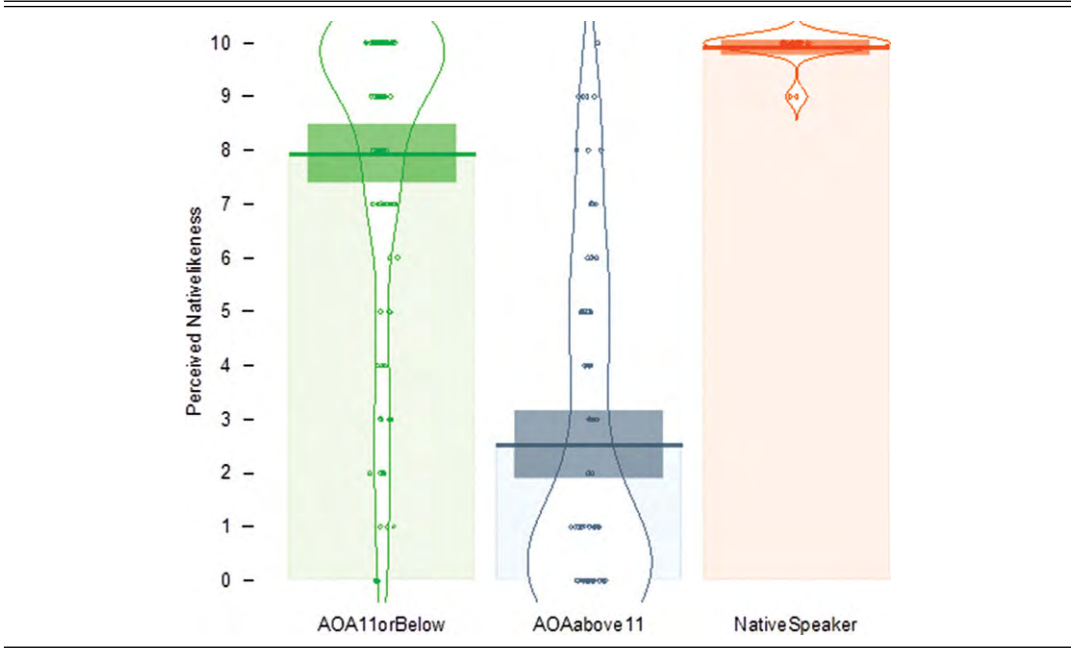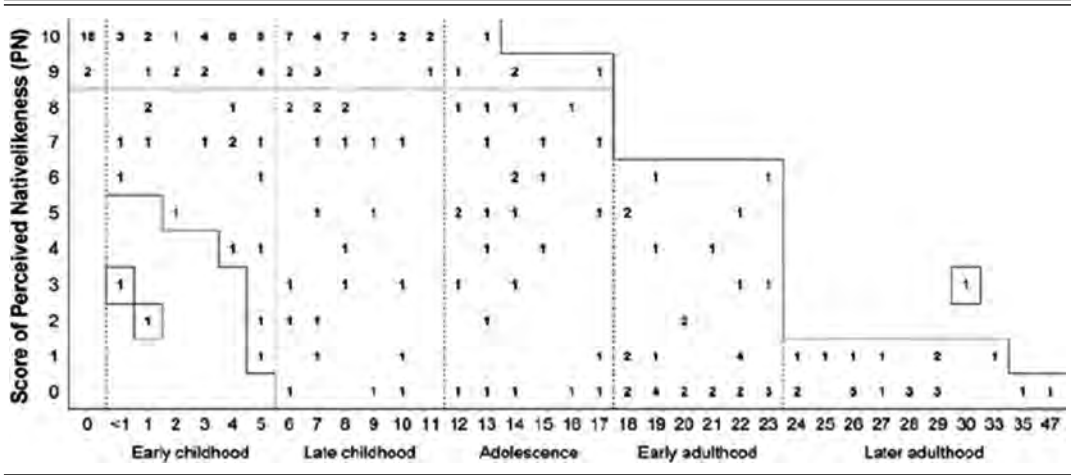

FIGURE 15

Abrahamsson & Hyltenstam's (2009) Plot of Judges' Perception of Nativeness for L2 Swedish Speakers, With Permission From Wiley

an English composition (Larson–Hall, 2015b; see Figure 16). The groups' sizes are essentially balanced, at $n = 17$ for the control group and $n = 16$ for the experimental (ALC Focus, where ALC refers to the alc.co.jp Web site) group. The black line imposed over the top shows the mean trend for each group. I created custom labels for the graphs by using the Paint program and also ad-

justed the graphs a bit to get them to line up with the same scale (although this was not perfect).

The parallel coordinate plots readily show that many more individual participants in the ALC focus group increased their number of corrections as compared to the control group, which is why the mean increases so much more steeply for the ALC focus group while it stays the same for the

FIGURE 16

Parallel Coordinate Plot of Larson–Hall (2015a) Data Showing Individual Changes From Pretest to Posttest
[Color figure can be viewed at wileyonlinelibrary.com]
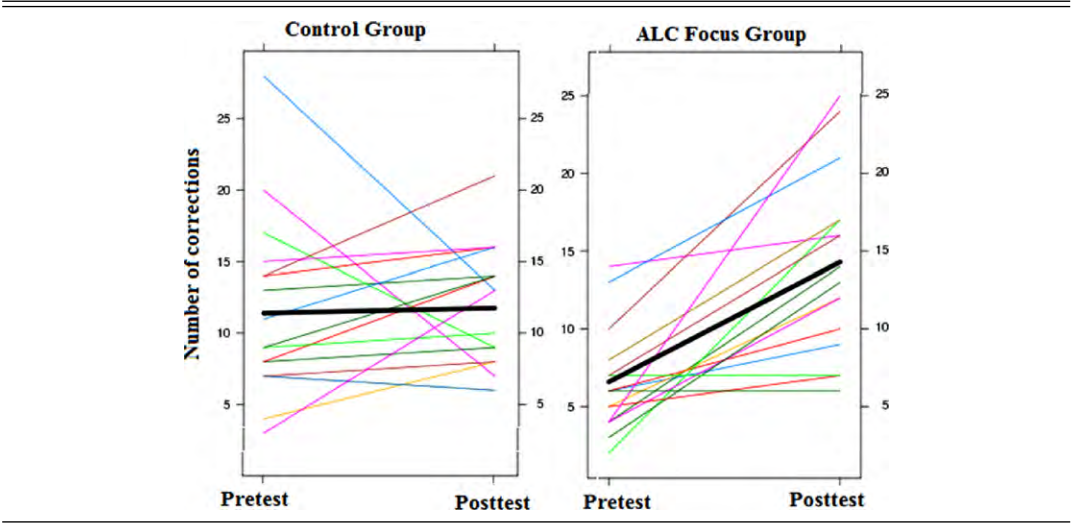
control group. It is also possible to see individual results, such that a few participants in the control group started out quite high, with many corrections on the pretest, but decreased markedly on the posttest, while that was not true of any of the participants in the ALC focus group. Readers could use the Paint program to add a line on top of the individual data at the mean levels in order to create a graphic with group as well as individual trends. Parallel coordinate plots with more than two groups can also be made.[9]

Jensen & Vinther (2003) contains an interesting graphic, which divides data into separate scatterplots by groups (see Figure 17). One might not generally think about using a scatterplot with group data (it is most often used for correlational data); but this graphic shows it is easy to create such a graphic when there is a pretest and posttest. The repetition of the scatterplot for the different groups gives the eye a way to compare groups quickly and also to see gains for each group on accurate phonological repetition in an elicited imitation task by plotting participant scores on the pretest on the x-axis and gains on the posttest on the y-axis. The vertical line at 537 is the overall mean pretest score, so the average gain for each group "can be read as the length of the vertical line between the regression line and the horizontal axis at point 537" (Jensen & Vinther, 2003, p. 400). Thus, the graphic shows that gains for the fsf listening condition (fast-slow-fast repetition of sentences) and the fss (fast-slow-slow) condition were certainly greater than for the

control group, who only performed the elicited imitation task but missed out on the eight training sessions. However, the regression line shows that gains were higher for students with lower pretest scores to begin with, and the fsf group seems to have had more students who scored high on the pretest so that gains on the posttest were not as impressive as for the fss group (since the negative slope of the line is steeper for the fsf group). This clever data accountable graph thus shows individual scores while still providing the overall picture of group results for different treatment conditions. This graphic is unusual but ultimately effective in the way it repeats the same design so that the eye is drawn to the comparison. It is essentially a small multiple.

Another way to examine different groups with paired data would be to put all of the data into one scatterplot while color coding the different groups and lines (the groups also use different plotting shapes), and this is shown for Jensen & Vinther's (2003) data in Figure 18. This kind of graphic, of course, would also work for correlational data. Here one can directly compare the mean gains (the length of the vertical line at point 537) and the steepness of the regression lines. I invite readers to consider which figure works better for comparing the information from the different groups. The R code for Figure 18 is found in the Appendix, and code for creating something like Figure 17 as I worked it out is also there (although not shown) in case readers might like to try the two graphs out with their own data.

FIGURE 17

Jensen & Vinther's (2003) Use of Multiple Scatterplots to Plot Pretest Scores Against Posttest Gains
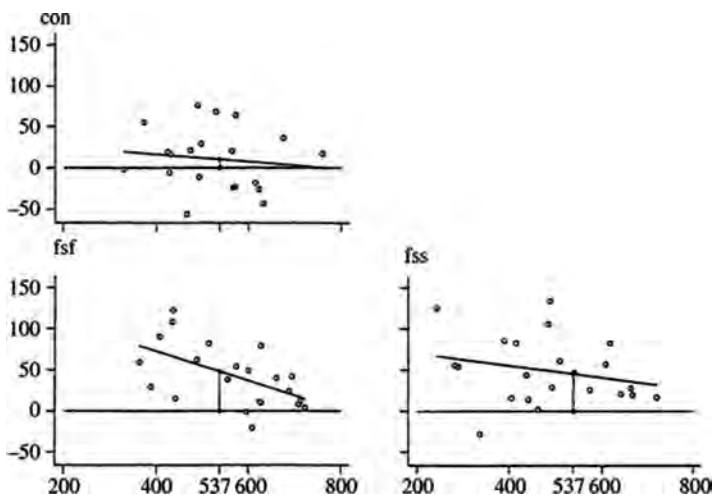
FIGURE 18

Jensen & Vinther's (2003) Data on One Scatterplot [Color figure can be viewed at wileyonlinelibrary.com]
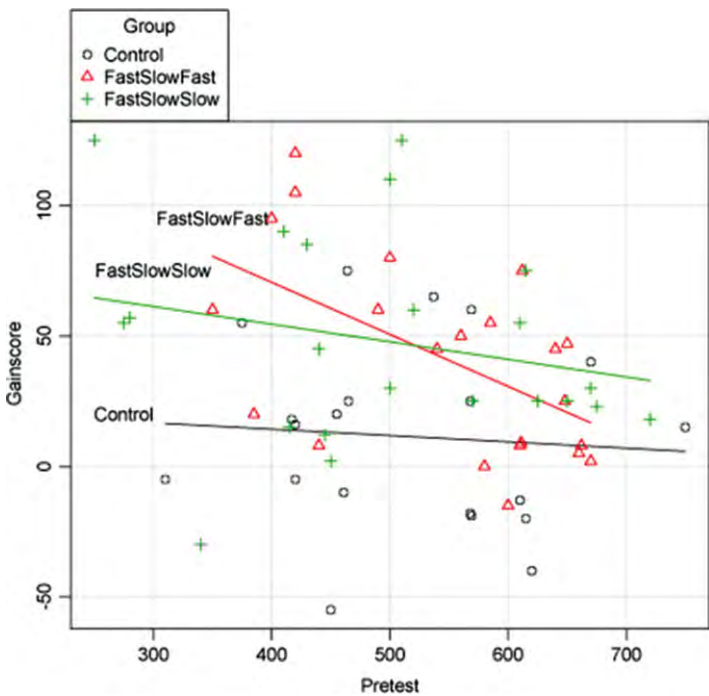
Graphics for Correlational Data. Scatterplots are unparalleled in their display of correlational data, and I cannot recommend any better graphics; however, I recommend that scatterplots contain both regression lines and Loess lines, as explained in Larson–Hall & Herrington (2009). Doing so allows readers to judge for themselves whether a straight line is the best approximation to the fit of the data; also, it is easy to create in any statistical program such as SPSS or R. The Loess

line may indicate that the data would be better modeled by a line with a bend in it or, alternatively, by splitting the data into separate groups on some variable.

Although scatterplots are wonderful data accountable graphics, they are currently much less popular in the L2 research literature than bar plots or line graphs. Table 1 at the beginning of the article showed that in the history of the three surveyed journals, scatterplots were used in less than 10% of the articles that contained graphics. In Hudson's (2015) survey of five journals over one year between 2012 and 2013, only 18 out of 157 instances were scatterplots (11%). Scatterplot matrices, which show scatterplots for a number of variables at one time, are even rarer; I did not encounter a single instance of such a graphic in my historical survey.

Nonstandardized Graphics. Another possibility for visualization are graphics customized to the format and needs of individual cases. Figure 15 showed this for the Abrahamsson & Hyltenstam (2009) data, which created a graphic where three factors could be displayed—total cumulative score on perceived nativeness by judges, number of participants achieving that score, and the age of onset of L2 study. Figure 19 introduces another data accountable graphic, this one created by Escudero and Boersma (2004) addressing the "Identification results of 30 Spanish listeners on the English /l/–/ɹ/ contrast [in each square, duration runs from 83 ms (left) to 176 ms (right), and F1 runs from 480 Hz (bottom) to 344 Hz (top)]." This unique graphic shows the results of listeners' judgments of synthetic vowels along a continuum that varied by duration and formant frequencies as /i/ or /ɪ/ (example: sheep or ship). Each box displays one individual's responses; the darker areas indicate a predominance of /i/ responses, while the lighter areas refer to a predominance of /ɪ/ responses. The solid line drawn shows the boundary where participants are equally likely to respond with /i/ as /ɪ/. This small multiple graphic is an effective way to display individual data while also putting the data together to look for group patterns: Once one understands the meaning of one box, one also understands the meaning of all the others. Thinking about the wealth of data behind these duration/frequency spectrograms, it is easy to appreciate just how much more informative a visual representation can be than raw numbers and how much further such visuals can go than summary data. Original designs like this need not take up much space, but do provide a large

impact on the reader for the space that they use.

Summary. This section has shown that all kinds of interesting graphics are possible and that thinking about what kind of graphic would best show the data and let the reader make visual comparisons is important. Data accountable and data rich graphics, which present the data in a way that lets the reader visually compare things are better methods for displaying and analyzing data than simple visual data summaries like the bar plot. As I hope to have argued successfully, they should be more convincing to the reader than an inferential statistical statement such as "There was a difference between groups, $t = 3.6$, $p = 0.001$."

The major problems with the bar plot and line graph for most data are that they provide very little information and waste space, are not data rich nor data accountable, and lack distributional information. Wilkinson and the APA task force (1999) claimed that graphics that do not show information about the shape of the distribution of the data can actually hinder accurate interpretation of the results of a study. To reiterate, the addition of error bars to a bar plot or line graph, which does possibly add inferential statistical information, can be problematic in terms of which kinds of error bars are being used: Most researchers are not able to use them to make quick judgments about statistical differences between groups. The last—and most obvious—point is simply that, beyond the bar plot, there are many other more informative or interesting designs that could be used to display data. In the recent past it may have been difficult to create such designs using available statistical software. But especially with the advent of the free statistical program, R, as well as online sites like langtest.jp, resources exist that allow researchers to fairly easily create very useful graphics for both exploratory and presentation purposes.

RECOMMENDATIONS

A field-wide strategy for improving data visualization must be multi-tiered and work from the bottom up as well as the top down. Papers and presentations by researchers in print and at conferences that include data accountable graphics can begin to influence others at the individual level to change the graphics that they choose. From the top down, I recommend that journal editors ask to see data accountable graphics in every

FIGURE 19

Escudero & Boersma's (2004) Graphic Showing Individual Trends as Well as Group Patterns
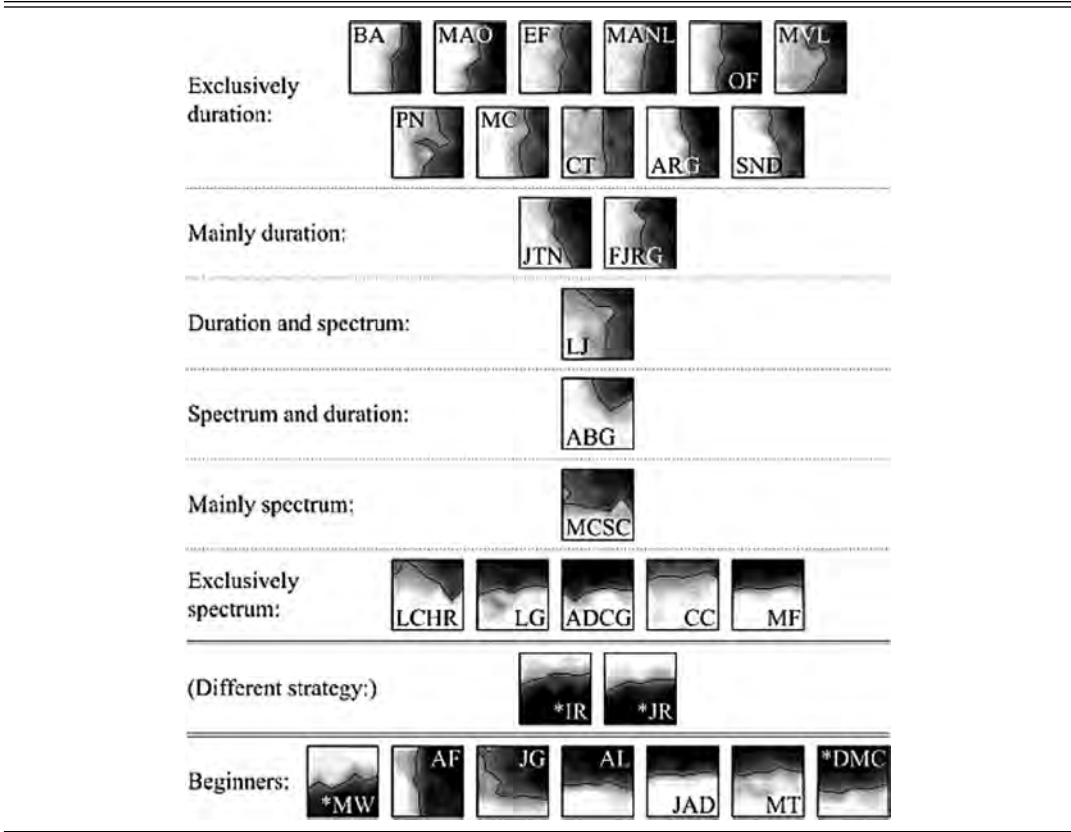
FIGURE 20

Graphic from the "BarBarPlots" Kickstarter Campaign [Color figure can be viewed at wileyonlinelibrary.com]
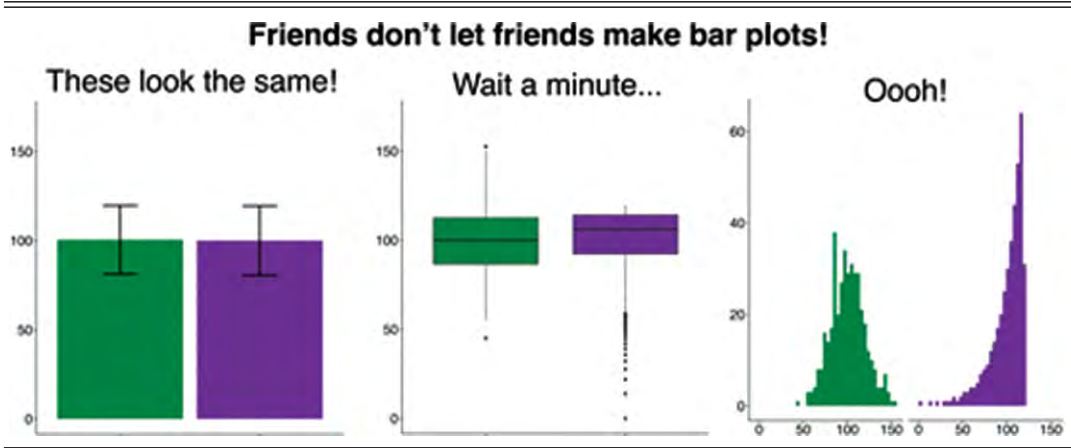
empirical article, whether those graphics get published or not. As an aside on this topic, a recent Kickstarter campaign called "Bar Bar Plots" aimed to send a message to major scientific journal editors in Psychology, Neuroscience, and Medicine that graphics should be improved to show clear and complete data visualization (see Figure 20; the funded campaign sent t-shirts with the logo in Figure 20 to journal editors). Reviewers can create a change in what is acceptable by commenting on

the graphics they see in submissions and asking for a data accountable graphic with a resubmission if one is not included. Graphics should be one of the topics addressed in L2 research methods courses. Future methodological handbooks in the field should continue to advocate for better graphics.

CONCLUSION

There are innumerable ways that L2 researchers can present their data in a visual form. Graphics should not be considered a frill in a research article, but instead an essential component in a quest to understand patterns and relationships among empirical variables. The very best ways to present data are usually the ones that present the most data and invite the reader to make comparisons among groups or testing times; these types of data will invariably be data rich, presenting many data points at a time. Data accountability in graphics will also enhance the ability of the reader to see both the forest and the trees in the data and to make comparisons among the residuals. This article has presented a wide range of graphs and information about how readers can make these graphs themselves. Several graphics, which are not readily available online or in the R program, have been illustrated with R code in the Appendix so readers could insert their own data and use these graphics. My hope is to have shown how interesting and exciting such graphics can be and how they can help move our work in language acquisition research forward as well.

NOTES

[1] The graphics in this paper will necessarily be in black and white in printed materials but are available in full color in the online version of this article.

[2] A reviewer asked in what cases it might be difficult to provide a presentation graphic. If the author had an institutional obligation not to reveal the data, it should still be possible to provide data rich graphics since they reveal distributions better than summary statistics but do not reveal individual data points. In the case of space considerations, online publication now provides ample space in appendices, if necessary, for graphics to be printed. In any case, they should be provided in manuscripts so that reviewers may see them, even if the editor declares there is not enough space in a publication for them. Last—but not least—Lane and Sandor (2009) stated, "There is an understandable desire on the part of researchers to show their data in a positive light. As a result, some may resist showing distributional data that reveal the variability and possible irregularities not apparent in a plot of means. However, this is clearly not a justifiable basis on which to omit distributional info" (pp. 241–242).

[3] See Plonsky, Egbert, & LaFlair (2015) for an account of the difficulty the authors had in trying to get raw data for a statistical review; there does seem to be a push to improve this situation—currently, *Language Learning* is participating in the Center for Open Science's push to encourage research transparency and data sharing by awarding science badges for researchers who share their data publicly (the 'Open Data' badge) as well as share their research materials and preregister their research designs. Recently the editorial boards of *MLJ* and *SSLA* have voted to join this initiative as well.

[4] In order to provide as much transparency as possible I wanted to place the data in a repository, and ISPCR seemed reputable and widely cited. As of this writing, the IRIS database (www.iris-database.org) only stores experimental materials and instruments, not raw data.

[5] Larson–Hall (2015a) provides guidance on creating beeswarm plots for three groups in Chapter 9 and an interaction plot for multiple groups and variables in Chapter 10 (both using the R program) that can explore the data, and both SPSS and R can create multiple histograms for different variables easily. SPSS instructions and R code for parallel coordinate plots can be found in Chapter 11 of Larson–Hall (2015a).

[6] Chapter 6 of Larson–Hall (2015a) gives directions for how to create such scatterplots in both SPSS and R.

[7] Chapter 7 in Larson–Hall (2015a) explains how to create scatterplot matrices in SPSS and R.

[8] In general, stem and leaf plots as well as histograms are rather unwieldy as presentation plots (Kampstra, 2008). On the next page is a stem and leaf of two groups with the same Abrahamsson & Hyltenstam (2009) data. I invite readers to judge for themselves how effective the stem and leaf plot is. The R code used to call it is given across the top of the figure.

```
>stem.leaf.backpack(ah2009$percnative[21:127], ah2009$percnative[128:215])
```

```
1 | 2: represents 1.2, leaf unit: 0.1
                         ah2009$percnative[21:127]                ah2009$percnative[128:215]

    3                                        000| 0* |00000000000000000000000000000000000   35
                                                | 0. |
    6                                        000| 1* |000000000000000000                     (18)
                                                | 1. |
   10                                       0000| 2* |000                                    37
                                                | 2. |
   14                                       0000| 3* |00000                                  34
                                                | 3. |
   17                                        000| 4* |0000                                   25
                                                | 4. |
   20                                        000| 5* |00000000                               25
                                                | 5. |
   22                                         00| 6* |00000                                  17
                                                | 6. |
   32                                 0000000000| 7* |000                                    12
                                                | 7. |
   41                                  000000001| 8* |0000                                   5
                                                | 8. |
   56                             000000000000001| 9* |0000                                  5
                                                | 9. |
   51  000000000000000000000000000000000000000001|10* |0                                     1

n:                     Age of Onset ≤ 11          107     88          Age of Onset ≥ 12
```

[9] Details for making these impressionistic (the numbers are not precise) parallel coordinate plots using R and SPSS can be found in Chapter 11 of Larson–Hall (2015a). However, I recently found myself needing to make a more precise parallel coordinate plot where the numbers for all of the measurements would align perfectly, and came up with a way to do it using the ggplot2 package in R. See the Appendix (under "Precise parallel coordinate plot") for the R code.

REFERENCES

Abrahamsson, N., & Hyltenstam, K. (2009). Age of onset and nativelikeness in a second language: Listener perception versus linguistic scrutiny. *Language Learning, 59*, 249–306.

Alves, D. (2013). *Tecnico*. Accessed 28 Jan 2016 at http://technical1293.blogspot.jp

Belia, S., Fidler, F., Williams, J., & Cumming, G. (2005). Researchers misunderstand confidence intervals and standard error bars. *Psychological Methods, 10*, 389–396.

Byrnes, H. (2013). Notes from the editor. *Modern Language Journal, 97*, 825–827.

Clark, W. H., & Clark, M. G. (1966). Achievement in elementary German under programmed and conventional instruction: A preliminary study. *Modern Language Journal 50*, 97–100.

Cleveland, W. S. (1985). *The elements of graphing data.* Monterey, CA: Wadsworth.

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*, 317–322.

Cumming, G. (2012). *Understanding the new statistics: Effect sizes, confidence intervals, and meta-analysis.* New York: Routledge/Taylor & Francis.

Cumming, G., Fidler, F., & Vaux, D. L. (2007). Error bars in experimental biology. *The Journal of Cell Biology, 177*, 7–11.

Dale, R., & Spivey, M. J. (2006). Unraveling the dyad: Using recurrence analysis to explore patterns of syntactic coordination between children and caregivers in conversation. *Language Learning, 56*, 391–430.

Ellis, N. C. (2015). Foreword. *Language Learning, 65*(S1), v–vi.

Escudero, P., & Boersma, P. (2004). Bridging the gap between L2 speech perception research and phonological theory. *Studies in Second Language Acquisition, 26*, 551–585.

Friendly, M. (2006). A brief history of data visualization. In C. Chen, W. H. Hardle, & A. Unwin (Eds.), *Handbook of computational statistics: Data visualization* (Vol. III, pp. 1–34). Berlin: Springer.

Graphical Abstracts. (2015). Accessed 28 Jan 2016 at https://www.elsevier.com/authors/journal-authors/graphical-abstract

Gullberg, M. (2006). Handling discourse: Gestures, reference tracking, and communication strategies in early L2. *Language Learning, 56*, 155–196.

Hudson, T. (2015). Presenting quantitative data visually. In L. Plonsky (Ed.), *Advancing quantitative methods in second language research* (pp. 78–105). New York: Routledge/Taylor & Francis.

Jensen, E. D., & Vinther, T. (2003). Exact repetition as input enhancement in second language acquisition. *Language Learning, 53*, 373–428.

Kampstra, P. (2008). Beanplot: A boxplot alternative for visual comparison of distributions. *Journal of Statistical Software, 28*, 1–9.

Lane, D. M., & Sandor, A. (2009). Designing better graphs by including distributional information and integrating words, numbers, and images. *Psychological Methods, 14*, 239–257.

Larkin, J. H., & Simon H. A. (1987). Why a diagram is (sometimes) worth 10,000 words. *Cognitive Science, 11*, 65–100.

Larson–Hall, J. (n.d.). Moving beyond the bar plot and line graph to create informative and attractive graphics. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor].

Larson–Hall, J. (2015a). *A guide to doing statistics in second language research using SPSS and R* (2nd ed.). New York: Routledge/Taylor & Francis.

Larson–Hall, J. (2015b). Using an online bilingual corpus dictionary to improve word choice and grammar. *Fukuoka JoGakuin Centre for the Study of English Language Teaching Journal* (英語教育研究センター), *3*, 99–109.

Larson–Hall, J., & Connell, S. (2005, October). *Evidence for a critical period: Remnants from forgotten languages.* Paper presented at the Second Language Research Forum, Teachers College, Columbia University, New York.

Larson–Hall, J., & Herrington, R. (2009). Improving data analysis in second language acquisition by utilizing modern developments in applied statistics. *Applied Linguistics, 31*, 368–390.

Larson–Hall, J., & Plonsky, L. (2015). Reporting and interpreting quantitative research findings: What gets reported and recommendations for the field. *Language Learning, 65*(S1), 127–159.

Lemper, L. L. (1925). Objective versus subjective tests in modern languages. *Modern Language Journal, 10*, 175–177.

Loewen, S., & Gass, S. (2009). The use of statistics in L2 acquisition research. *Language Teaching, 42*, 181–196.

Norris, J. M., Ross, S. J., & Schoonen, R. (2015). Improving second language quantitative research. *Language Learning, 65*(S1), 1–8.

Pearson, B. Z., & Fernandez, S. C. (1994). Patterns of interaction in the lexical growth in two languages of bilingual infants and toddlers. *Language Learning, 44*, 617–653.

Phillips, N. D. (2016). The pirate plot (2.0)—the RDI plotting choice of R pirates. Accessed 30 June 2016 at http://nathanieldphillips.com/2016/04/pirateplot-2-0-the-rdi-plotting-choice-of-r-pirates/

Plonsky, L., Egbert, J., & LaFlair, G. T. (2015). Bootstrapping in Applied Linguistics: Assessing its potential using shared data. *Applied Linguistics, 36*, 591–610.

Poehner, M. E., & Lantolf, J. P. (2013). Bringing the ZPD into the equation: Capturing L2 development during computerized dynamic assessment (C-DA). *Language Teaching Research, 17*, 323–342.

Saito, K. (2013). The acquisitional value of recasts in instructed second language speech learning: Teaching the perception and production of English /ɹ/ to adult Japanese learners. *Language Learning, 63*, 499–529.

Sandor, A., & Lane, D. (2007). *Graph use in psychology journals.* Poster presented at the annual conference of the Houston Chapter of the Human Factors and Ergonomics Society, Houston, TX.

Sommers, M. S., & Barcroft, J. (2013). Effects of referent token variability on L2 vocabulary learning. *Language Learning, 63*, 186–210.

Tufte, E. (1990). *Envisioning information.* Cheshire, CT: Graphics Press.

Tufte, E. (1997). *Visual explanations: Images and quantities, evidence and narrative.* Cheshire, CT: Graphics Press.

Tufte, E. (2001). *The visual display of quantitative information.* Cheshire, CT: Graphics Press.

Tufte, E. (2006). *Beautiful evidence.* Cheshire, CT: Graphics Press.

Tufte, E. (2014, January 16). *Sparkline theory and practice Edward Tufte* [Online forum comment]. Accessed 28 Jan 2016 at http://www.edwardtufte.com/bboard/q-and-a-fetch-msg?msg_id=0001OR

Tukey, J. W. (1977). *Exploratory data analysis.* Menlo Park, CA: Addison–Wesley.

Unwin, A. (2008). Good graphics? In C.–H. Chen, W. Hardle, & A. Unwin (Eds.), *Handbook of data visualization* (pp. 57–78). Berlin: Springer.

Unwin, A., Chen, C.–H., & Hardle, W. (2008). Introduction. In C.–H. Chen, W. Hardle, & A. Unwin (Eds.), *Handbook of data visualization* (pp. 3–12). Berlin: Springer.

Wainer, H. (1996). Depicting error. *American Statistician, 50*, 101–111.

Ware, C. (2008). *Visual thinking for design.* Amsterdam: Elsevier.

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. doi:10.1371/journal.pbio.1002128

Wilkinson, L., & The Task Force on Statistical Inference (1999). Statistical methods in psychology journals: Guidelines and explanations. *American Psychologist, 54*, 594–604.

APPENDIX

R Code
FIGURE 9: Clark and Clark mini histograms

```
#First, enter data. I used the original graphic for the data but used the lowest number
#as my entry in each case (for example, for Listening Comprehension there were 3
#cases of the count "13-15" and I entered 3 "13"s)
C1←c(16, 16, 19, 19, 19, 19, 19, 22, 22, 22, 25, 25, 25, 25, 25, 28, 28, 28, 31, 31)
P1←c(13, 13, 13, 16, 16, 16, 16, 16, 16, 19, 19, 22, 22, 22, 22, 28, 31, 34, 37, 37)
C2←c(26, 26, 26, 30, 30, 34, 34, 34, 34, 38, 38, 38, 38, 42, 42, 42, 42, 42,46, 50)
P2←c(34, 38, 38, 38, 38, 38, 42, 42, 42, 42, 42, 46, 46, 46, 50, 54, 58, 58, 58, 66)
C3←c(13, 16, 16, 19, 22, 22, 22, 25, 25, 28, 31, 31, 31, 31, 34, 37, 37, 37, 40, 43)
P3←c(10, 13, 13, 13, 16, 16, 16, 16, 16, 16, 19, 19, 19, 19, 25, 31, 31, 37, 40, 43)
C4←c(39, 39, 44, 49, 49, 49, 49, 49, 54, 54, 59, 64, 64, 64, 64, 64, 69, 69, 74, 84)
P4←c(29, 34, 34, 34, 39, 39, 39, 44, 44, 44, 49, 49, 54, 54, 54, 64, 69, 69, 69, 84)

library(plotrix) #Use command install.packages("plotrix") if needed first
par(mfrow = c(4, 1))

mh1 ← multhist(Group1, xlab = " ", ylab = "Frequency", main = " Listening Comprehension",
col = c("black", "light grey"),breaks = seq(12, 39, by = 3),
names.arg = c("13-15", "16-18",
"19-21", "22-24", "25-27", "28-30", "31-33", "34-36", "37-39"))
box(bty = "l", col = "black") #this draws a line on the x-axis
legend.text←c("Group C","Group P")
legend(locator(1), legend = legend.text, col = c("black", "light grey"), pch = 15, bty = "n", cex = 0.8)
#You need to click on the graphic to insert the legend before running the next
#part of the command (and so on after each graphic)

mh2 ← multhist(Group2, xlab = " ", ylab = "Frequency", main = "Speaking",
col = c("black", "light grey"),breaks = seq(25, 69, by = 4),
names.arg = c("26-29", "30-33",
"34-37", "38-41", "42-45", "46-49", "50-53", "54-57", "58-61", "62-65", "66-69"))
box(bty = "l", col = "black")
legend.text←c("Group C","Group P")
legend(locator(1), legend = legend.text, col = c("black", "light grey"), pch = 15, bty = "n", cex = 0.8)

mh3 ← multhist(Group3, xlab = " ", ylab = "Frequency", main = "Reading",
col = c("black", "light grey"),breaks = seq(9, 45, by = 3),
names.arg = c("10-12", "13-15",
"16-18", "19-21", "22-24", "25-27", "28-30", "31-33", "34-36", "37-39", "40-42", "43-45"))
box(bty = "l", col = "black")
legend.text←c("Group C","Group P")
legend(locator(1), legend = legend.text, col = c("black", "light grey"), pch = 15, bty = "n", cex = 0.8)

mh4 ← multhist(Group4, xlab = " ", ylab = "Frequency", main = " Writing ",
col = c("black", "light grey"),breaks = seq(28, 88, by = 5),
names.arg = c("29-33", "34-38", "39-43", "44-48", "49-53", "54-58", "59-63", "64-68", "60-73", "74-78", "79-
83", "84-88"))
box(bty = "l", col = "black")
legend.text←c("Group C","Group P")
legend(locator(1), legend = legend.text, col = c("black", "light grey"), pch = 15, bty = "n", cex = 0.8)
```

FIGURE 13: Overlaid boxplots for three groups

```
#The data, called ah2009, can be manually entered by looking at the data in Figure 15
#of this paper
#Alternatively, download the SPSS datafile from the website:
#http://www.routledgetextbooks.com/textbooks/9781138024571/spss_data.php
y←ah2009$percnative[129:215]
x←ah2009$percnative[22:128]

library(beeswarm) #use install.packages("beeswarm") if needed

#my data was appearing in the wrong order so I created a factor
#then added it to the end of my dataframe (giving new name)
OrderedGroup = factor(rep(1:3, c(20,107,88)))
levels(OrderedGroup) = c("NS", "AOA 11 or below", "AOA above 11")
ah2009Ordered←cbind(ah2009,Ordered Group)
#so this orders it so that NS comes first, then 11 or Under then 11 and Above
boxplot(percnative~OrderedGroup, data = ah2009Ordered,
outline = F, xlab = "Means and +/-1 SDs are displayed in red.", ylab = "Score of Perceived Nativelike-
ness")

# use outline = F to avoid double-plotting outliers, if any
beeswarm(percnative~OrderedGroup,data = ah2009Ordered, cex = .6,
#I had to reduce size of plotting points to get separation
#of the groups for some lines of the plot
Col = 4, pch = 16, add = T)

x←ah2009Ordered$percnative[1:20]
#I just manually noted the line numbers
#of the data that was in Group 1 (native speakers)
y←ah2009Ordered$percnative[21:127] #Group 2 (11 and below) data
z←ah2009Ordered$percnative[128:215] #Group 3 (above 11) data

points(1.2, mean(x), pch = 18, col = "red", cex = 2)
arrows(1.2, mean(x), 1.2, mean(x) + sd(x), length = 0.1, angle = 45, col = "red")
arrows(1.2, mean(x), 1.2, mean(x) - sd(x), length = 0.1, angle = 45, col = "red")
points(2.2, mean(y), pch = 18, col = "red", cex = 2)
arrows(2.2, mean(y), 2.2, mean(y) + sd(y), length = 0.1, angle = 45, col = "red")
arrows(2.2, mean(y), 2.2, mean(y) - sd(y), length = 0.1, angle = 45, col = "red")
points(3.2, mean(z), pch = 18, col = "red", cex = 2)
arrows(3.2, mean(z), 3.2, mean(z) + sd(z), length = 0.1, angle = 45, col = "red")
arrows(3.2, mean(z), 3.2, mean(z) - sd(z), length = 0.1, angle = 45, col = " red ")
```

FIGURE 14: Pirate plot

```
library(yarrr)
#Use the ah2009 data; see Figure 13 instructions
pirateplot(formula = percnative~group,data = ah2009,xlab = " ", ylab = "Perceived Nativelikeness",
main = " ",pal = "appletv", point.cex = .5, point.pch = 1,point.o = 1, inf.o = .5, line.o = 1, bean.o = 1,
inf = "ci")
```

FIGURE 17: R code to approximate Jensen & Vinther's (2003) figure

```
#dataset estimated from Jensen & Vinther (2003) data
Group = factor(rep(1:3, c(20, 22, 22)))
Pretest = c(310, 375, 417, 420, 420, 450, 455, 461, 464, 465, 537, 568, 569, 568, 569, 610, 615, 620, 670,
750, 350, 385, 400, 420, 420, 440, 490, 500, 540, 560, 580, 585, 600,610, 611, 612, 640, 648, 650, 660, 662,
670, 250, 275, 280, 340, 410, 415, 430, 440, 445, 450, 500, 500, 510, 520, 570, 610, 615, 625, 650, 670, 675,
720)
```

```
Gainscore = c(-5, 55, 18, 16, -5, -55, 20, -10, 75, 25, 65, -18, -19, 25, 60, -13, -20, -40, 40, 15, 60, 20, 95, 105,
120, 8, 60, 80, 45, 50, 0, 55, -15, 8, 9, 75, 45, 25, 47, 5, 8, 2, 125, 55, 57, -30, 90, 15, 85, 45, 12, 2,30, 110,
125, 60, 25, 55, 75, 25, 25, 30, 23, 18)
jv2003←data.frame (Group, Pretest, Gainscore) #but this is matrix not dataframe
jv2003$Group←as.factor(jv2003$Group) #give levels names now
levels(jv2003$Group) = c("Control", "FastSlowFast", "FastSlowSlow")


par(mfrow = c(2,2))


plot(Gainscore[1:20]~Pretest[1:20],
xlab = " ", ylab = " ", ylim = range(c(-50, 150)), xlim = range(c(200,800)), xaxt = "n", main = "Control",
data = jv2003)
abline(0,0)
reg←lm(jv2003$Gainscore[1:20]~jv2003$ Pretest[1:20])
abline(reg)
axis(1, at = c(200,400,537,600,800))
segments(537,0,537,11)


plot(Gainscore[21:42]~Pretest[21:42],
xlab = "Pretest score", ylab = "Gain score", main = "FSF", ylim = range(c(-50, 150)), xlim =
range(c(200,800)), xaxt = "n", data = jv2003)
abline(0,0)
reg←lm(jv2003$Gainscore[21:42]~jv2003$ Pretest[21:42])
abline(reg)
axis(1, at = c(200,400,537,600,800))
segments(537,0,537,43)


plot(Gainscore[43:64]~Pretest[43:64],
xlab = " ", ylab = " ", main = "FSS", ylim = range(c(-50, 150)), xlim = range(c(200,800)), xaxt = "n",
data = jv2003)
abline(0,0)
reg←lm(jv2003$Gainscore[43:64]~jv2003$ Pretest[43:64])
abline(reg)
axis(1, at = c(200,400,537,600,800))
segments(537,0,537,45)
```

FIGURE 18: Jensen & Vinther scatterplot with 3 groups on one plot
#Use jv2003 data input for Figure 17

```
library(car) #Use install.packages("car") first if necessary


scatterplot(Gainscore~Pretest, by.groups = T, groups = Group, smoother = F, reg.line = lm,
xlim = range(c(200,800)), data = jv2003)
text(350,95, c("FastSlowFast"))
text(250, 70, c("FastSlowSlow"))
text(275,20, c("Control"))
abline(0,0)
axis(1,at = c(537))
segments(537,0,537,45)
text(500, 40, c("FSS = 43"), col = "green")
text(600, 40, c("FSF = 45"), col = "red")
text(580, 15, c("Con = 11"))
```

PRECISE PARALLEL COORDINATE PLOT

```
library(plyr)#Use install.packages("plyr") if necessary, first
library(ggplot2)#Also install this first if not installed
#Enter raw data
Student←c("A01", "A02", "A03", "A04", "A05", "A06", "A07", "A08", "A09", "A10", "A11", "A12", "A13",
"A14", "A15", "A01", "A02", "A03", "A04", "A05", "A06", "A07", "A08", "A09", "A11", "A12", "A13", "A14",
"A15","A01", "A02", "A03", "A04", "A05", "A06", "A07", "A08", "A09", "A12", "A13", "A14", "A15","A01",
"A02", "A03", "A04","A06", "A07", "A08", "A09", "A12", "A13", "A14", "A15","A01", "A02", "A03", "A04",
"A05", "A06", "A07", "A08", "A09", "A10", "A11", "A12", "A13", "A14", "A15")
Percentage.Correct ←c(56, 61, 54, 47, 82, 36, 31, 39, 61, 35, 47, 42, 95, 81, 93, 35, 40, 23, 20,68, 6, 46,
19, 30, 12, 39, 64, 39, 89, 43, 42, 12, 13, 38, 6, 21, 14, 20, 44, 23, 17, 77, 10, 19, 2, 4, 2, 10, 2, 6, 24, 19, 1,
25, 5, 20, 2, 3, 9, 3, 6, 4, 6, 1, 1, 19, 26, 2, 23)
Time.in.Weeks←rep(c(0, 3.5, 7, 49, 90), c(15, 14, 13, 12, 15))
Bierling = data.frame(Student,Percentage. Correct,Time.in.Weeks)


ggplot(Bierling, aes(Time.in.Weeks, Percentage.Correct, group = Student)) +
geom_line(aes(color = Student)) + #gives each individual a differ ent color line
geom text(aes(label = Student), size = 2) + #puts in a label for each individual at
#every node on the x-axis
theme(legend.position = "none") +
#deletes the legend for each individual student
scale_x_continuous(breaks = c(0, 3.5, 7, 49, 90)) +
# inserts x-axis ticks proportional to the number of weeks
ggtitle("Production Data") #overall title
```