Association for Psychological Science

Perspectives on Psychological Science
2016, Vol. 11(5) 692–701

DOI: 10.1177/1745691616663875
pps.sagepub.com
SAGE

Graphical Descriptives: A Way to Improve Data Transparency and Methodological Rigor in Psychology

Louis Tay[1], Scott Parrigon[1], Qiming Huang[1], and James M. LeBreton[2]
[1]Purdue University and [2]Pennsylvania State University

Abstract

Several calls have recently been issued to the social sciences for enhanced transparency of research processes and enhanced rigor in the methodological treatment of data and data analytics. We propose the use of *graphical descriptives* (GDs) as one mechanism for responding to both of these calls. GDs provide a way to visually examine data. They serve as quick and efficient tools for checking data distributions, variable relations, and the potential appropriateness of different statistical analyses (e.g., do data meet the minimum assumptions for a particular analytic method). Consequently, we believe that GDs can promote increased transparency in the journal review process, encourage best practices for data analysis, and promote a more inductive approach to understanding psychological data. We illustrate the value of potentially including GDs as a step in the peer-review process and provide a user-friendly online resource (www.graphicaldescriptives.org) for researchers interested in including data visualizations in their research. We conclude with suggestions on how GDs can be expanded and developed to enhance transparency.

Keywords

data transparency, data visualization, graphics, best practices

Our field of psychology—and social science in general—is recognizing that greater transparency is needed in the research process (Open Science Collaboration, 2012). To this end, some have proposed providing infrastructure to document and share study materials (Open Science Framework), calling for greater transparency in data collection (Simmons, Nelson, & Simonsohn, 2011), encouraging greater rigor and transparency in the reporting of statistical analyses (Cortina, 2015), and the promotion of analytic strategies that focus on estimation based on effect sizes and confidence intervals (Cumming, 2014). The goals of these proposals are to improve openness and methodological rigor, thereby promoting scientific reproducibility for psychological research, which has recently been called into question.

In line with these goals, we propose that *graphical descriptives* (GDs), or the visualization of research data, can be used in the research process and eventually be developed into a routine component of the publication process. GDs can serve as quick and efficient checks for authors, reviewers, and the general scientific audience to assess data distributions, variable relations, outliers, and the appropriateness of statistical analyses while maintaining a level of information privacy and security.

In order to promote scientific reproducibility, GDs have several core features. The first feature is to optimally display data; that is, to provide as much information about the data as possible while balancing visual efficacy (e.g., reducing clutter) and the potential need for data privacy so that proprietary data sets cannot be recreated from the visual presentation. The second feature is to provide efficient, visual checks of data analytic assumptions and inferences. The final feature is a more practical one—GDs can be conveniently applied and reported to a wide array of data, with minimal burden on researchers, reviewers, editors, or publishers. To this end, we propose the use of GDs in the research and publication

Corresponding Author:
Louis Tay, Department of Psychological Sciences, Purdue University, 703 Third Street, West Lafayette, IN 47907
E-mail: stay@purdue.edu

process and offer an initial set of online tools for researchers interested in including the results of data visualizations in their future manuscripts. Our long-term goal is to expand this set of tools to accommodate a wide range of designs and data so that it can eventually be incorporated into the research and publication process.

Why GDs?

Ensuring data fidelity and the requisite assumptions for statistical analyses are vital for improving methodological rigor in the psychological sciences. One way to promote high standards is via the public disclosure of all data, which is practiced in some fields (e.g., medicine; Godlee & Groves, 2012; Mello et al., 2013).[1] Yet, such a tactic may be limited because review and publication delays incurred on journals may create resistance to such an implementation. Further, the time costs for reviewers and/or the additional financial costs for statisticians to analyze the data may also limit the effectiveness of such an approach. In addition, some fields of psychology (e.g., organizational psychology, educational psychology, and health psychology) often rely on proprietary data, which cannot be easily shared. Finally, opening access to data (whilst certainly serving to increase data transparency) may be viewed as reciprocally unfair by researchers who spend substantial amounts of time, energy, and money collecting data, only to be required to share those data with researchers who have not invested similar levels of time, energy, and financial resources into the data collection process. As such, despite the ideal of public disclosure of data, there may be practical restrictions on its implementation.

Given these practical restrictions, alternatives to the public disclosure of data have been suggested and include encouraging researchers to submit the data set used in their analyses to the journal when they submit their manuscript for peer review or, similarly, to submit the materials used for their procedures. Other calls have encouraged researchers to preregister their design and analytic plans and to utilize the new statistics reporting methods that focus on estimation of effect sizes, confidence intervals, and meta-analyses (e.g., Cumming, 2014; Funder et al., 2014; Nosek & Lakens, 2014).

The proposed GDs approach complements these other approaches by utilizing the strengths of data visualizations to convey rich, nuanced information not easily gleaned by more traditional methods of data reporting (e.g., correlations, means, variances). GDs provide insights into fundamental data issues including assessing basic data distributions, (non-)linearity between variables, outliers, patterns of missing data, and the tenability of assumptions underlying various statistical analyses. Although such topics are regularly covered in courses on psychological methods and statistics, it is less common to see such topics explicated in our research. We examined the journals *Psychological Science*, *Journal of Personality and Social Psychology*, *Journal of Applied Psychology*, and *Journal of Abnormal Psychology* over the past 5 years (2010–2014) and found that the base rates associated with reporting fundamental data issues ranged from 5% to 17%. As shown in Figure 1, in spite of the increasing focus on methodological rigor in the field, there does not appear to be evidence of any positive trends in reporting data checks.

Part of the reason for the lack of reporting of data checks may stem from journal space limitations. However, the growing acceptance of online supplemental material can provide the necessary space for its inclusion.

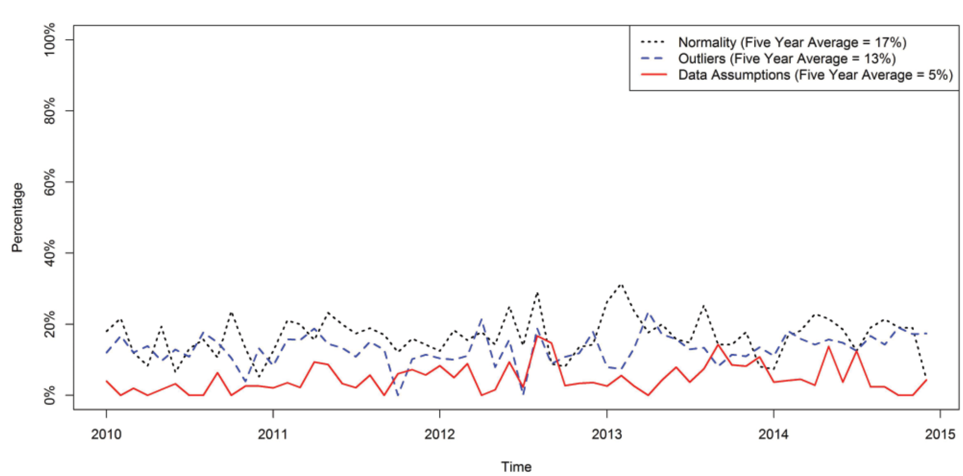

Fig. 1. Percentage of studies in a sample of major journals (*Psychological Science*, *Journal of Personality and Social Psychology*, *Journal of Abnormal Psychology,* and *Journal of Applied Psychology*) reporting data checks over 5 years (2010–2014).

Moreover, the proliferation and availability of new data visualization tools, in traditional statistical software as well as that offered in the free software R (R Development Core Team, 2015), allows convenient data visualizations generalizable for a variety of data types and analyses (Chang, 2013; Murrell, 2011). However, these tools are not widely utilized by psychologists due to a lack of exposure, lack of fluency with R, and/or a lack of pressure from reviewers and journal editors to include GDs of data.

We propose that the use of GDs have several distinct advantages:

- GDs can provide information about univariate and bivariate data distributions, which may improve our understanding of the sample characteristics (mode of the distribution, skew, kurtosis, minimum/maximum value, etc.) and the accompanying psychological phenomena. Such visualizations yield a richer integrative picture of data as compared to what is typically gleaned by a cursory review of the descriptive statistics and correlations, although these are important in themselves to examine as well. Because information presented visually is often more engaging than numeric presentation and because visualization simultaneously captures multiple phenomena (e.g., mode, skew, kurtosis), methodological issues may be easier to identify with visualization than with a table of numbers.
- The incorporation of GDs can promote best practices for researchers to check their data prior to submission for publication. For example, a basic "checklist" (cf. "Journal Article Reporting Standards") asking authors to verify that they explicitly examined the extent to which their data meet the critical assumptions underlying the statistical tools utilized provides an efficient, additional layer of checks in the peer-review system. Thus, GDs can be used to facilitate open discussion concerning the extent to which data satisfied the requisite assumptions of different statistical analyses (e.g., Brandt, 2012; Ullrich & Schluter, 2012).
- By including GDs in the peer-review process whenever possible, we can strengthen the quality of research by further increasing the transparency of data while maintaining a level of information privacy with minimal burden on researchers, reviewers, and editors. Importantly, GDs could help increase transparency not only in situations where data cannot be shared (e.g., for proprietary reasons), but also when used as a supplement to data sharing because they can provide a quicker, easier-to-digest snapshot of various data issues than a raw data set is often able to provide.
- GDs can enable possible detection of errors in data coding or influential outliers. Positively, GDs can reveal important but unexplored trends that were not of immediate interest to the primary researchers. This can lead to multiple data uses for different research projects and increased opportunities for scientific collaborations.
- The incorporation of GDs can also increase the use of inductive methods within psychology (Locke, 2007). The growing interest in "Big Data" is expected to foment the increased use of data visualizations, and GDs are expected to become an increasingly integral tool for uncovering new and interesting psychological phenomenon (Keim, Kohlhammer, Ellis, & Mansmann, 2010).

Given these advantages, we propose that the use of GDs can enable the field to positively improve data transparency and methodological rigor. In the next section, we illustrate the different types of GDs that can be easily incorporated as part of most research. In Table 1, we present pertinent information to take note of when interpreting each type of GD.

Illustrating Types of GDs

General descriptives

Although it is common to report descriptive statistics (means, modes, standard deviations, and correlations), GDs can be used to visualize univariate distributions and bivariate relations, which can provide substantially more information at a glance. This is not to suggest that GDs should replace common descriptive statistics—rather, they supplement what is commonly reported as well as what is less commonly reported (minimum score, maximum score, skewness, kurtosis). There are also some aspects, such as unimodality, which are easier to identify through visualization. We illustrate the usefulness of GDs with an example data set (N = 123) that we collected on Amazon's Mechanical Turk (MTurk). This data set includes the self-reported variables of age, gender, race, ACT scores, college GPA, Conscientiousness (α = .92; Goldberg, 1999), and Intellectual Openness (α = .92; Woo et al., 2014). Table 2 contains the descriptive statistics that are routinely included in most journal submissions.

As shown in Figure 2, GDs allow us to examine the univariate distributions of the data (as can be seen by the histograms and kernel density estimates displayed across the diagonal) in order to assess whether variables have single or multiple modes, the extent to which they might

Table 1. Recommended Information to Consider When Interpreting Graphical Descriptives

Type	Issues/questions to consider
General descriptives	— Are the minimum/maximum values within expected range? — Are the univariate distributions unimodal or multimodal? — Are there large skews in the univariate distributions? — Are there outliers or extreme values in the univariate or bivariate plots? — Are the relations between variables linear or nonlinear (curvilinear or piecewise linear—i.e., composed of different straight-line functions at different intervals)? — Are there other abnormalities in the data (miscoded values, errors in entry, range restriction, etc.)?
Group mean differences	— Are there outliers, extreme values, or additional subgroups that could lead to incorrect inferences when making comparisons between two observed groups? — Are the distributions fairly unimodal?
Moderator analyses	— Are there outliers or extreme values driving or attenuating predicted moderation effects? — It is assumed that a linear relationship holds between the independent and dependent variable across groups. Is this the case? — To what extent are the grouping variable and the predictor variable related? A strong relationship may imply non-overlapping predictor variable distributions that could indicate nonlinearity between the independent variable and the dependent variable as an alternate explanation.
Statistical assumptions	— Given the assumption of linearity, does the relation between the fitted values and the residuals exhibit a curved or otherwise nonlinear pattern? — Is there a constant spread of the residuals at all levels of the fitted values or is there a systematic restriction or expansion (e.g., a "shotgun" effect) across the fitted values? — Do the residuals approximate a normal distribution? — Are there any outlying residual values that might be unduly influencing the results?

be skewed, and the potential presence of outliers (Cohen, Cohen, West, & Aiken, 2003). A cursory glance shows that the age, ACT scores, and college GPA variables are all fairly skewed. The uneven distribution in age is likely because MTurk samples are more demographically diverse than the typical undergraduate psychology student population (Buhrmester, Kwang, & Gosling, 2011), and GPA scores are known to be relatively skewed (e.g., Chansky, 1964). Importantly, there were a number of individuals who appear to be outliers based upon their low ACT scores. An examination of the ACT scores reveals that a number of participants had entered zero, with one participant entering a score of three. The former is certainly an invalid score (as zero is outside of the possible range of ACT scores: 1–36) and the latter is likely an invalid or typographical error (as individuals with such low scores would likely have difficulty navigating the Internet and completing the MTurk registration process). Moreover, these invalid data points artificially inflate the correlation between ACT scores and Intellectual Openness from $r = .24$ (omitting these observations) to $r = .35$ (including these observations), an error that we would likely overlook if we were to simply examine the typically reported means, standard deviations, and correlations.

Figure 2 also displays the bivariate relations through scatterplots and fitted local regression (loess) lines (the panel below the diagonal). We highlight one observation here. The relation between Conscientiousness and Intellectual Openness may be nonlinear, such that individuals who are above average on Conscientiousness show strong positive relations with Intellectual Openness

Table 2. Correlations and Descriptive Statistics for a Mechanical Turk Sample

Variable	*M*	*SD*	1	2	3	4	5	6
1. Age	24.41	6.16	—					
2. ACT Total	25.80	7.78	−.18	—				
3. College GPA	3.42	.45	.02	.21*	—			
4. Conscientiousness	3.50	.66	.07	.15	.24*	(.92)		
5. Intellectual Openness	3.77	.54	.14	.35*	.11	.42*	(.92)	
6. Gender	.50	.50	−.13	−.15	−.17	−.11	−.21*	—

Note: $N = 123$
*$p < .05$.

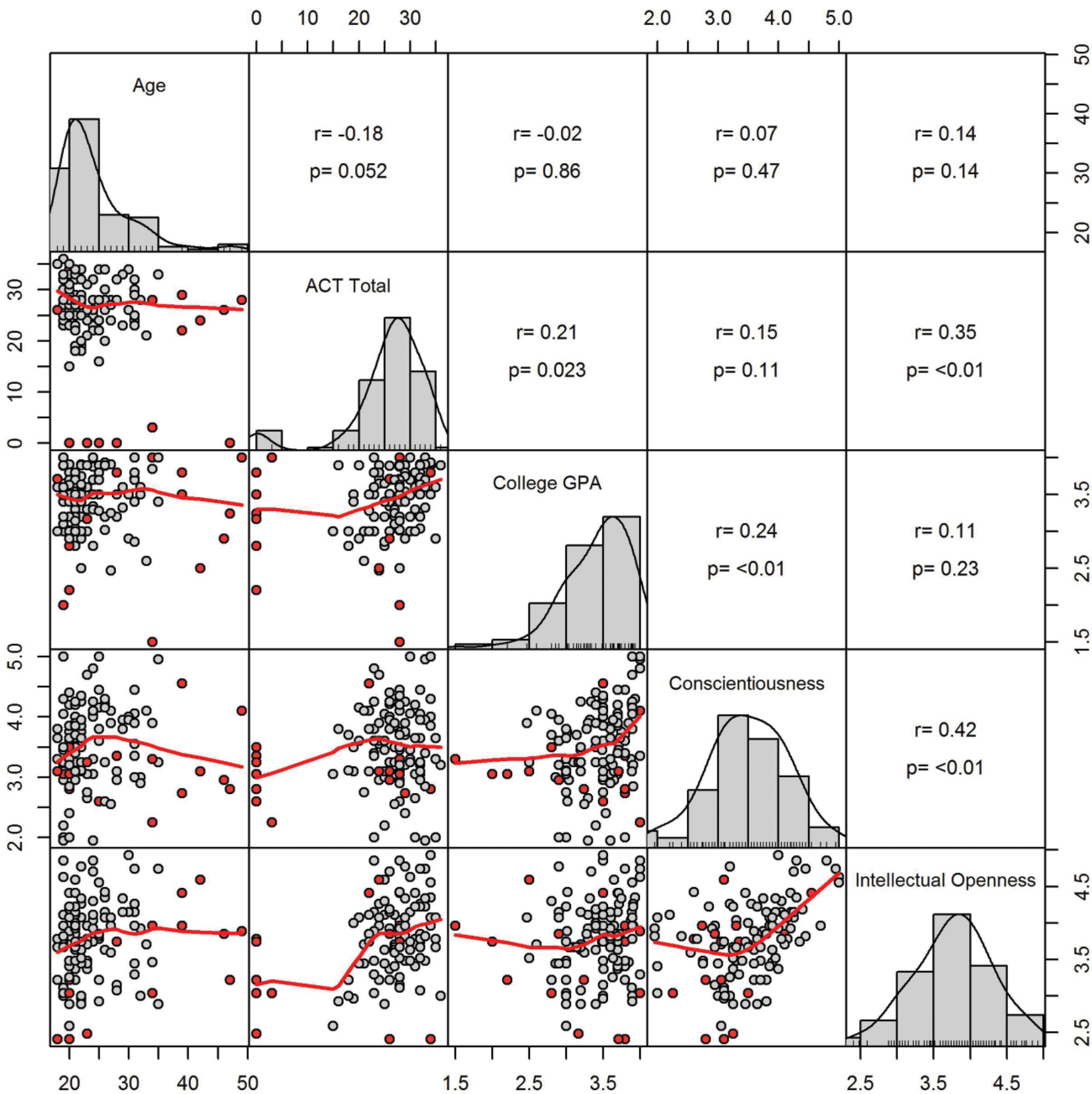

Fig. 2. Descriptives: Scatterplots and univariate distributions. Red dots are used to highlight potential univariate outliers (as indicated by 1.5 times the interquartile range).

($r = .59$, $p < .001$), whereas individuals lower than average on Conscientiousness show no such relation ($r = .08$, $p = .56$). This does not appear to be a function of outliers. As such, one may venture a threshold hypothesis whereby past a minimal level of Conscientiousness, intellectual engagement rises as Conscientiousness increases. However, without visualizing these bivariate distributions, this nonlinear trend would likely go undetected and researchers could erroneously conclude that any statistical assumptions of linearity between these variables were met and that a consistent, moderate relation between Conscientiousness and Intellectual Openness was present ($r = .42$, $p < .05$).

Group mean differences

Assessing and reporting group differences is an indispensable part of psychological research. The current standard of reporting effect sizes, though helpful in interpreting findings beyond statistical significance (American Psychological Association, 2009), does not reveal if specific data points and/or subgroups drive or obscure the

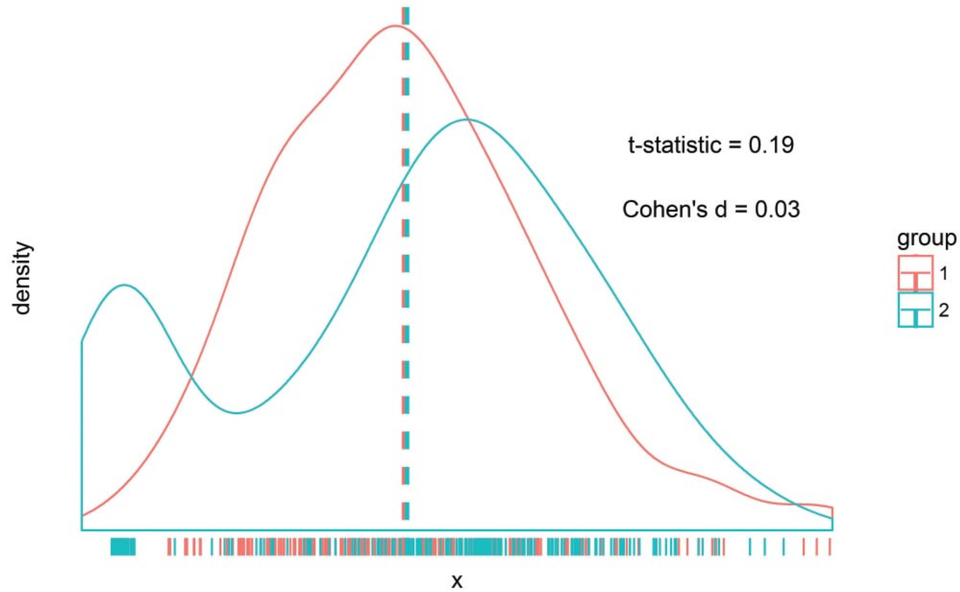

Fig. 3. Mean difference: Density and rug plots.

effects. This may be especially important when smaller sample sizes are used (previous reviews suggest median *N*s < 200 in psychological research; Marszalek, Barber, Kohlhart, & Holmes, 2011; Shen et al., 2011) and group means—which are typically used—can be vulnerable to outliers. The use of GDs can serve to elucidate potentially influential observations when conducting comparisons between group means. As illustrated in Figure 3 with two groups (ns = 100) an initial mean difference analysis does not reveal significant differences (t = .19, p > .05, d = 0.03) between the two groups. However, an examination of the density and rug plot reveals that this may be a consequence of two subgroups in Group 2, where there are a subgroup of individuals on the low end ($m < -2.0$) and another subgroup of individuals who are slightly higher ($m \approx .40$) than Group 1. A mean difference GD can provide a more nuanced interpretation of possible drivers of these effects (be they statistically significant or statistically nonsignificant), effects that might go undetected using traditional reporting standards. Importantly, these effects should be cross-validated in another sample to prevent capitalizing on chance.

Moderator analyses

Another analysis that is commonly of interest to researchers is the exploration of moderating effects, where researchers seek to determine if the strength or direction of effect between two variables (e.g., y and x) changes contingent on another variable (e.g., z). We show how GDs, in terms of scatterplots and data distributions, can be more informative than simply graphing the interaction lines, with y, x, and z, typically denoted as the outcome, predictor, and moderator variable, respectively. As shown in Figure 4, there appear to be disordinal (i.e., crossover) effects based on the interaction lines; that is, the direction of the effect changes as a function of the moderating variable z. The moderator here is a grouping variable z (e.g., experimental condition, gender, high–low of a median split). Only examining the interaction lines commonly leads to the conclusion that the moderator variable largely accounts for the direction of effects. However, by using GDs, the additional scatterplots and data distributions in Figure 4 reveal that this may also be accounted for by a curvilinear relation (i.e., inverted-U relations) between the predictor and outcome variable. This is because the data distributions of the x variable corresponds with the low and high values of the moderator, suggesting that an alternative or supplemental explanation may be that the direction of the effect changes as x increases (Cortina, 1993). This moderating effect found with only the interaction lines could limit theoretical explanation and insight. As a hypothetical example, finding that gender moderates the direction of the effect between mathematics self-efficacy (x) and performance (y) could further be understood (or primarily explained) by the observation that females have lower mathematics self-efficacy than males.

Statistical assumptions

GDs can also be used to establish the tenability of statistical assumptions in order to ensure accurate statistical inferences (Cohen et al., 2003). We illustrate the use of

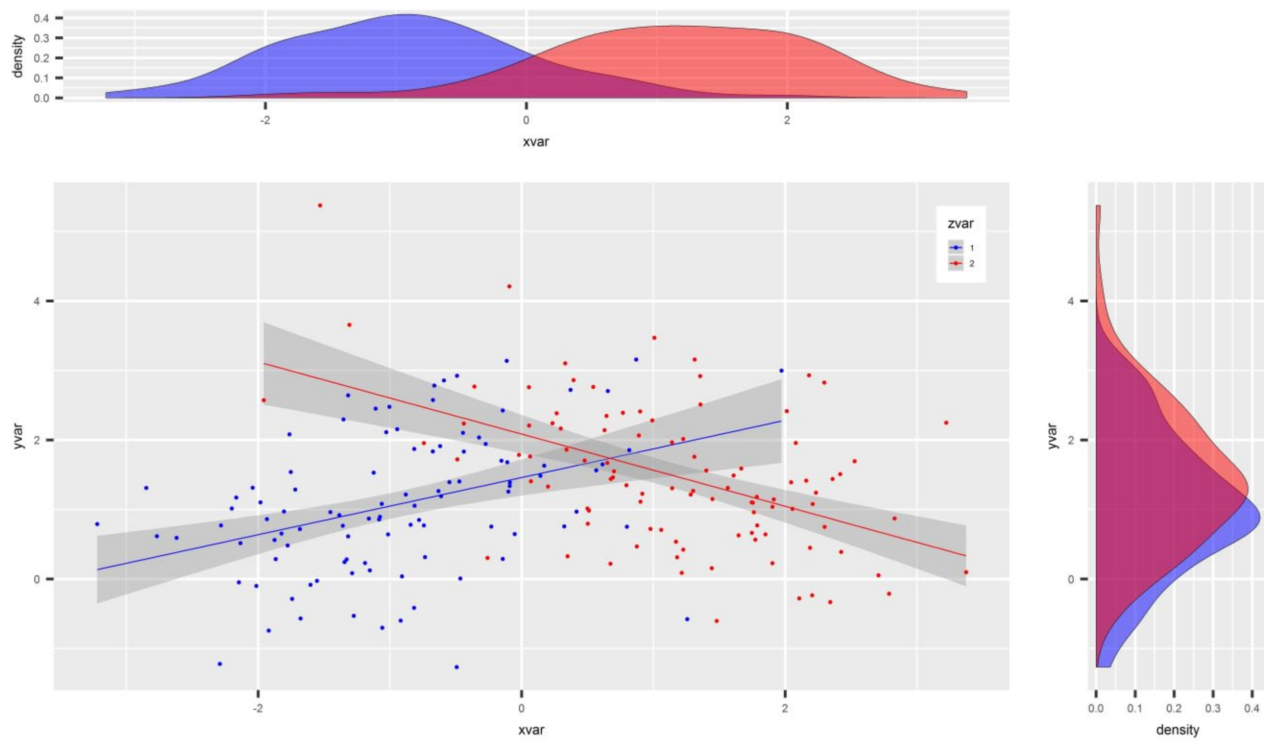

Fig. 4. Plot of moderation analyses.

GDs in examining the statistical assumptions of regression. For this illustration, we used a subset of data from the Gallup-Purdue Index[2] (n = 17,703) to predict student loan amounts. The variables used for this analysis are displayed in Table 3. The unstandardized regression results, reveal that only year graduated (β = .333, p < .05) significantly predicted loan amount, such that the more recently an individual graduated, the larger the amount of student loans that were taken out.

To assess the extent to which the data actually support such a conclusion, we can use statistical assumptions GDs, which can be help assess the tenability of the assumptions underlying this analysis. We examine several assumptions of linear regression here: (a) correctly specifying the form of the relationship between the independent and dependent variables, (b) constant error variance (homoscedasticity), and (c) the normality of residuals.

Table 3. Regressions Predicting Loan Amount for the Gallup-Purdue Sample

	Loans	Square root loans[a]
Variable	β	β
Year Graduated	.333*	.341*
Marital Status	.006	.014*
Gender	−.026	−.037*

Note: n = 17,703.
[a]Results using heteroscedasticity-consistent standard errors are reported.
*p < .05

GDs for these three statistical assumptions are provided Figure 5. The first assumption—the form of the relationship between the independent and dependent variables—can be examined using the top-left plot ("Residuals vs Fitted" in Fig. 5). We expect the red loess line in this plot to be horizontal, which appears to generally be the case, suggesting that the form of the relations might be correctly specified. The second assumption—that there is homoscedasticity or constant error variance—can be evaluated using this same plot. If the variance (the height spanned by the plotted black points) is substantially restricted and/or expanded at various points along the fitted values, then this assumption of constant error variance may be violated. The plot indicates that the variance of the residuals seems to be expanding from left to right (i.e., variance increases when fitted values for the dependent variable are higher), providing evidence for the lack of tenability of this assumption. A second way to evaluate this assumption can be found in the plot in the middle row on the left side of Figure 5 ("Scale Location"), which depicts the fitted values against the square-root of the absolute standardized

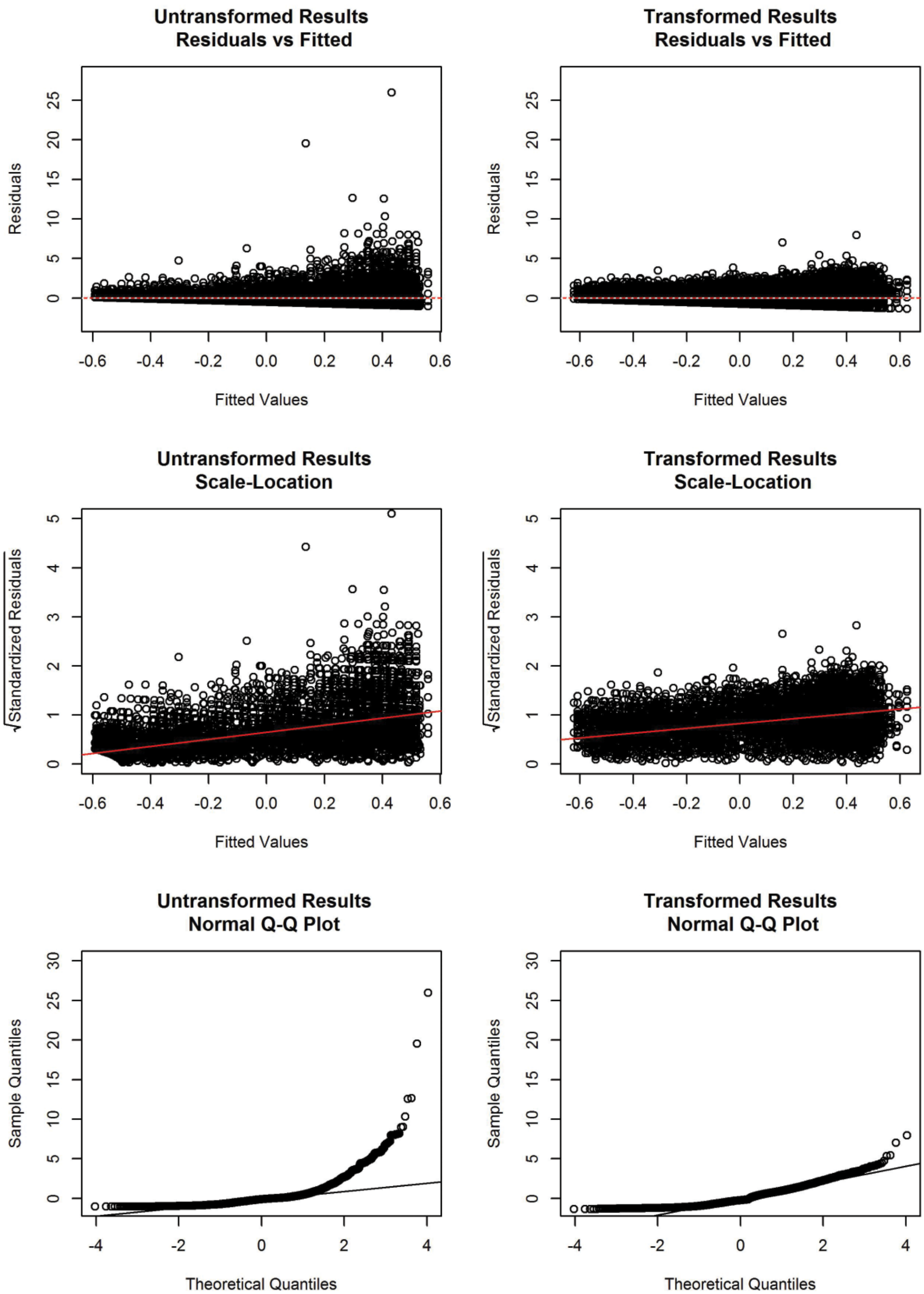

Fig. 5. Graphical descriptives for linear regression diagnostics.

residuals. This plot rescales and "folds" the top-left plot along the zero-point of the residuals (the values below the zero value in the top-left plot are now positioned at an equivalent, positive value at the same fitted value), thus providing a slightly different view to assess homoscedasticity. The fit line in this plot is increasing, as opposed to showing a flat slope, indicating that the variance of the residuals across levels of the fitted values is increasing (i.e., error variance does not seem to be constant). Finally, for the assumption of normality of residuals, we use the bottom-left plot ("Normal Q-Q Plot" in Fig. 5). The black line in this plot represents the values expected if the residuals conformed to a normal distribution. As can be seen, many of the data points fall above the line that would be expected for a normal distribution, suggesting the violation of this assumption. Thus, these plots generally suggest that the assumptions of linear regression are not supported. Further visual inspection of the dependent variable (loans) indicates that the variable is highly skewed, suggesting that this is likely a major contributing factor to the assumption violations that we see here. To address this problem, we applied a square root transformation to the loans outcome. Analyses using this transformed outcome revealed better conformity to assumptions. The updated assumptions plots (shown on the right side of Fig. 5) show that the normality of the residuals (the bottom-right plot) has been much improved by this transformation (with the residuals more accurately falling onto the dotted line signifying a normal distribution) and that the other plots are also somewhat improved. However, we would still be somewhat concerned about heteroscedasticity within the data after the transformation—notice that there still seems to be some evidence of increasing variance (i.e., some evidence of an increasing spread in the heights of the black dots in the top-right plot and evidence for a positive trend of the red loess line in the middle-right plot). These observations tell us that we would be wise to account for possible heteroscedasticity in our analyses. To address this, we can conduct our analyses with heteroscedasticity-consistent standard errors (Long & Ervin, 2000). The updated analyses show that in addition to year graduated, both marital status (b = .014, $p < .05$) and gender (b = −.037, $p < .05$) were also significant predictors of student loans (see Table 3). Individuals who are married typically have higher student loans than individuals who are single, and females typically have fewer student loans than their male counterparts. This demonstration shows that, even with a very large sample size (n = 17,703), a failure to use GDs to assess the assumptions of the analyses conducted may lead researchers to erroneous conclusions and missed findings.

Discussion

Transparency and methodological rigor are foundational for science and there are multiple means of promoting these values. In this article, we propose that GDs serve as a way to complement other important methods and incentives that are being proposed and established. We have attempted to illustrate how various instantiations of GDs can further promote transparency and methodological rigor by revealing characteristics of data that can prompt further examination, induction, and collaboration. In addition, GDs may serve as an initial quality control check for problematic data and to confirm that basic statistical assumptions have been met and increase the trustworthiness of our inferences and conclusions.

We propose that by encouraging GDs as a standard in our field, we can enhance methodological transparency and rigor. One possible path to increasing the use of GDs is to include GDs as supplemental materials in journal submissions. Another possibility is to provide digital signatures for all submissions noting that data analysis checks have been completed prior to manuscript submission. This will also help to promote common standards for data assumption checks that may be shared by authors, reviewers, and the general scientific audience. To provide additional incentives, the current use of Open Practices Badges in APS could be augmented to include badges for the use of GDs.

In this vein, we have developed a working website (www.graphicaldescriptives.org) that enables researchers to generate GDs. In its current form, our website emphasizes the more rudimentary plots as illustrated in our article. At this juncture, such plots are likely most useful for data that do not have sophisticated data structures (e.g., dyadic data, nested data). Despite that, we believe that for now the GD plots such as scatter plots and mean difference plots are still useful to a large number of research projects.

Moving forward, we will expand the tools available on the website so that it may be used to generate GDs for different types of data formats and structures. This would include, not exhaustively, experimental data with different designs, multilevel data, and time series data. In the future, we also plan to expand the website to include interactive tools that facilitate researchers uploading their data to a secure website so that other researchers can interactively visualize data and, if interested, obtain permissions to collaborate on extensions with the primary authors, especially for sensitive or proprietary data. With the growing interest in Big Data, a critical challenge for the current platform would be to provide GD tools that would be able to handle the processing demands of Big

Data sets and remaining visually efficacious by utilizing techniques such as data binning and automatic scaling of plots. Finally, we are collaborating with data visualization experts to implement methods for promoting data transparency while also ensuring privacy protections so that proprietary/confidential data sets cannot be recovered wholesale from data visualizations.

To conclude, we believe that there will be a growing need for GDs in the field of psychology if we are to increase data transparency and methodological rigor. We advocate for the use of GDs in the research and publication process. We hope that key stakeholders and researchers can join in our efforts to further the use of GDs in journals and the development of the GDs platform.

Declaration of Conflicting Interests

The authors declared that they had no conflicts of interest with respect to their authorship or the publication of this article.

Notes

1. We note here that, where possible, federally funded projects are also obligated to provide public access to digital data from scientific publications.
2. Information: http://products.gallup.com/168857/gallup-purdue-index-inaugural-national-report.aspx; Methodology: http://www.gallup.com/174167/temp-methodology-gallup-purdue-index-methodology.aspx

References

American Psychological Association. (2009). *Publication manual* (6th ed.). Washington, DC: American Psychological Association.

Brandt, M. (2012). Nasty data can still be real: A reply to Ullrich and Schluter. *Psychological Science*, *23*, 826–827.

Buhrmester, M., Kwang, T., & Gosling, S. D. (2011). Amazon's Mechanical Turk: A new source of inexpensive, yet high-quality, data? *Perspectives on Psychological Science*, *6*, 3–5.

Chang, W. (2013). *R graphics cookbook*. Sebastopol, CA: O'Reilly Media.

Chansky, N. (1964). Progress of promoted and repeating grade I failures. *The Journal of Experimental Education*, *32*, 225–237.

Cleveland, W. S., & Devlin, S. J. (1988). Locally weighted regression: an approach to regression analysis by local fitting. *Journal of the American Statistical Association*, *83*, 596–610.

Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2003). *Applied multiple regression/correlation analysis for the behavioral sciences* (3rd ed.). Mahwah, NJ: Erlbaum.

Cortina, J. M. (1993). What is coefficient alpha? An examination of theory and applications. *Journal of Applied Psychology*, *78*, 98–104.

Cortina, J. M. (2015). *Presidential address*. Speech presented at the Society for Industrial and Organizational Psychology, Philadelphia, PA.

Cumming, G. (2014). The new statistics: Why and how. *Psychological Science*, *25*, 7–29.

Funder, D. C., Levine, J. M., Mackie, D. M., Morf, C. C., Vazire, S., & West, S. G. (2014). Improving the dependability of research in personality and social psychology recommendations for research and educational practice. *Personality and Social Psychology Review*, *18*, 3–12.

Godlee, F., & Groves, T. (2012). The new BMJ policy on sharing data from drug and device trials. *British Medical Journal*, *345*, e7888.

Goldberg, L. R. (1999). A broad-bandwidth, public domain, personality inventory measuring the lower-level facets of several five-factor models. *Personality Psychology in Europe*, *7*, 7–28.

Keim, D., Kohlhammer, J., Ellis, G., & Mansmann, F. (Eds.). (2010). *Mastering the information age: Solving problems with visual analytics*. Goslar, Germany: Eurographics Association.

Locke, E. A. (2007). The case for inductive theory building? *Journal of Management*, *33*, 867–890.

Long, J. S., & Ervin, L. H. (2000). Using heteroscedasticity consistent standard errors in the linear regression model. *The American Statistician*, *54*, 217–224.

Marszalek, J. M., Barber, C., Kohlhart, J., & Holmes, C. B. (2011). Sample size in psychological research over the past 30 years. *Perceptual Motor Skills*, *112*, 331–348.

Mello, M. M., Francer, J. K., Wilenzick, M., Teden, P., Bierer, B. E., & Barnes, M. (2013). Preparing for responsible sharing of clinical trial data. *New England Journal of Medicine*, *369*, 1651–1658.

Murrell, P. (2011). *R graphics* (2nd ed.). London, United Kingdom: Chapman & Hall.

Nosek, B. A., & Lakens, D. (2014). Registered reports. *Social Psychology*, *45*, 137–141.

Open Science Collaboration. (2012). An open, large-scale collaborative effort to estimate the reproducibility of psychological science. *Perspectives on Psychological Science*, *7*, 657–660.

R Development Core Team. (2015). R: A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from http://www.R-project.org

Shen, W., Kiger, T. B., Davies, S. E., Rasch, R. L., Simon, K. M., & Ones, D. S. (2011). Samples in applied psychology: Over a decade of research in review. *Journal of Applied Psychology*, *96*, 1055–1064.

Simmons, J. P., Nelson, L. D., & Simonsohn, U. (2011). False-positive psychology: Undisclosed flexibility in data collection and analysis allows presenting anything as significant. *Psychological Science*, *22*, 1359–1366.

Ullrich, J., & Schluter, E. (2012). Detecting nasty data with simple plots of complex models: Comment on Brandt (2011). *Psychological Science*, *23*, 824–825.

Woo, S. E., Chernyshenko, O. S., Longley, A., Zhang, Z. X., Chiu, C. Y., & Stark, S. E. (2014). Openness to experience: Its lower level structure, measurement, and cross-cultural equivalence. *Journal of Personality Assessment*, *96*, 29–45.