

Marquette University

**e-Publications@Marquette**

---

Psychology Faculty Research and Publications

Psychology, Department of

---

Summer 2021

## **Remote Teaching of Publication-Quality, Single-Case Graphs in Microsoft Excel**

Robert K. Lehardy

Kevin C. Luczynski

Stephanie Hood

Ciobha A. McKeown

Follow this and additional works at: [https://epublications.marquette.edu/psych\\_fac](https://epublications.marquette.edu/psych_fac)



Part of the [Psychology Commons](#)

---

Marquette University

e-Publications@Marquette

***Psychology Faculty Research and Publications/College of Arts and Sciences***

***This paper is NOT THE PUBLISHED VERSION.***

Access the published version via the link in the citation below.

*Journal of Applied Behavior Analysis*, Vol. 54, No. 3 (Summer 2021): 1265-1280. [DOI](#). This article is © Wiley and permission has been granted for this version to appear in [e-Publications@Marquette](#). Wiley does not grant permission for this article to be further copied/distributed or hosted elsewhere without the express permission from Wiley.

# Remote Teaching of Publication-Quality, Single-Case Graphs in Microsoft Excel

Robert K. Lehardy

University of Nebraska Medical Center's Munroe-Meyer Institute, Omaha, NE

Kevin C. Luczynski

University of Nebraska Medical Center's Munroe-Meyer Institute, Omaha, NE

Stephanie A. Hood

Department of Psychology, Marquette University, Milwaukee, WI

Ciobha A. McKeown

Department of Psychology, University of Florida, Gainesville, FL

## Abstract

Microsoft Excel is ubiquitous, cost-effective, and can be used to create publication-quality single-case design graphs. We systematically replicated the GraphPad Prism video tutorial by Mitteer et al. (2018) to teach 24 master's students to create multiple-baseline graphs using Excel 2016. Students' mean accuracy on the multiple-baseline graph was 25% in pretraining, 86% with the video tutorial, and 96% with the review checklist. Next, students used the same video tutorial to create multielement and reversal graphs. Students' mean accuracy on the multielement graph was 93% with video tutorial and 94% with review checklist, and accuracy on the reversal

graph was 82% with video tutorial and 94% with review checklist. Students reported moderate to high satisfaction with both training components. The results support scientist-practitioners using the video tutorial and review checklists to create three common graphs using Excel 2016, Excel 2019, and Excel Office 365.

The graphical display of data is prevalent in journals for medicine, physics, chemistry, biology, engineering, and behavior analysis (i.e., *Journal of Applied Behavior Analysis* and *Journal of the Experimental Analysis of Behavior*), indicating that graph use is a primary trait of scientific disciplines (Arsenault et al., **2006**; Kubina et al., **2008**). Graph use is a hallmark of behavior analysis because graphs enable researchers and practitioners to explore, analyze, summarize, and communicate data in a clear, precise, and efficient way (Tuftte, **2001**). Furthermore, creating and updating graphs is a core competency to obtain certification for all behavior analysis practitioners (Behavior Analysis Certification Board, **2012, 2018**).

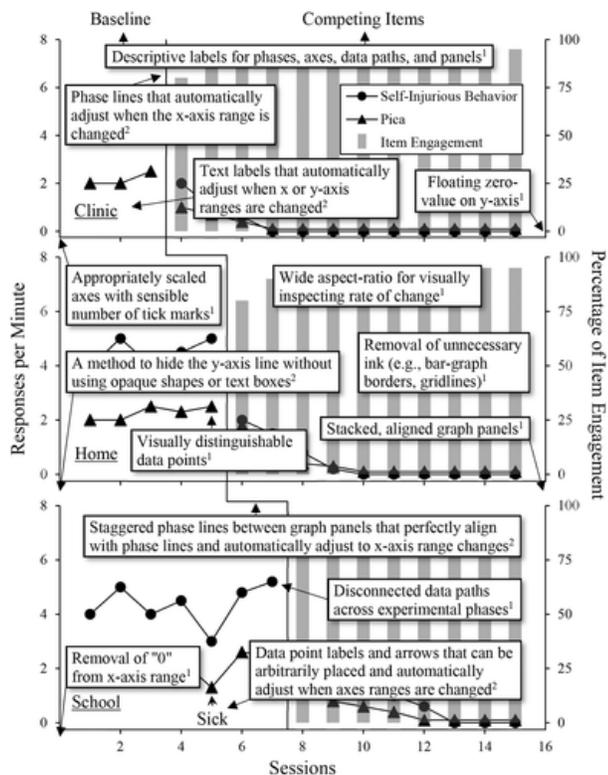
To enhance visual inspection, Mitteer et al. (**2018**) reviewed books, journal guidelines, and publications to identify functional and aesthetic (i.e., publication-quality) features of single-case graphs. Next, the authors developed a video tutorial to teach four participants, who were Registered Behavior Technicians™, to create a multiple-baseline graph with publication-quality features using GraphPad Prism. All participants created the target graph with 85% or higher accuracy. Three participants used the video tutorial to create additional multiple-baseline graphs with 90% or higher accuracy, which maintained at high levels after one month. The outcomes are notable because incorporating publication-quality features may improve visual inspection of level, trend, and variability of data during clinical or peer review. The outcomes also contribute to a body of literature that has demonstrated the efficacy of video tutorials to teach graphing skills (Tyner & Fienup, **2015**; Vanselow & Bourret, **2012**). Video tutorials are likely efficacious because participants can learn by listening to instructions while observing a step-by-step demonstration for each part of the graph. For example, in Tyner and Fienup (**2015**), students who used a video tutorial created graphs more accurately and quickly than students who used a written and pictorial tutorial. Video tutorials can also be widely disseminated to individuals through websites like YouTube. For example, the Prism video tutorial by Mitteer et al. has over 800 unique views. One video tutorial to create a reversal graph using Microsoft Excel has more than 34,000 unique views.

Mitteer et al. (**2018**) chose Prism over Excel because it could be more easily used to create graphs with publication-quality features. The authors acknowledged Excel's lower cost and in-cell calculation abilities as strengths but noted that graphs created using Excel often lacked publication-quality features. Due to Excel's ubiquitous use, with a worldwide adoption rate in over 71% of educational institutions and 56% of all institutions (Bitglass, **2018**), it would be worthwhile to create and validate the efficacy of a video tutorial to teach individuals to create multiple-baseline graphs with publication-quality features using Excel.

In addition, the tutorial should describe how to create phase lines, phase lines across panels (i.e., staircase lines), text labels, and arrows as automated features. Automated features remain in position when the axes' ranges change as new data are added or the graph size is changed, which is advantageous when graphs are updated over time. Researchers have created written and pictorial tutorials to describe how to insert automated phase lines (Dubuque, **2015**; Fuller & Dubuque, **2019**), an automated method to hide the y-axis line after floating the y-axis without using opaque shapes (Deochand et al., **2015**), and automated text labels (Deochand, **2017**). However, researchers have not described methods to insert staircase lines as an automated feature, and no video tutorial exists to teach individuals to create publication-quality graphs with automated features in Excel 2016 or later versions.

It was notable that Mitteer et al. (**2018**) created multiple-baseline graphs with multiple data paths and phase lines because these features are necessary when creating multielement and reversal graphs. For this reason, the authors suggested that their video tutorial for a multiple-baseline graph could be used to create multielement or reversal graphs, but they did not conduct this generality test.

In summary, creating single-case design graphs is a critical skill for behavior analysts, there are publication-quality features that are advantageous to incorporate into graphs, and researchers have published evidence that video tutorials are efficacious in teaching graphing skills. If individuals have access to Excel and do not prefer to buy or learn how to use Prism, Excel can be used to create graphs with publication-quality and automated features. The first aim of this study was to systematically replicate the procedures by Mitteer et al. (2018) and test the efficacy of a video tutorial on creating multiple-baseline graphs with publication-quality and automated features using Excel 2016 (see Figure 1 for a depiction of all features). The second aim of this study was to determine if individuals could use the video tutorial for the multiple-baseline graph to create multielement and reversal graphs without a graph-specific video tutorial.



**Figure 1** Publication-Quality Features of Graphs

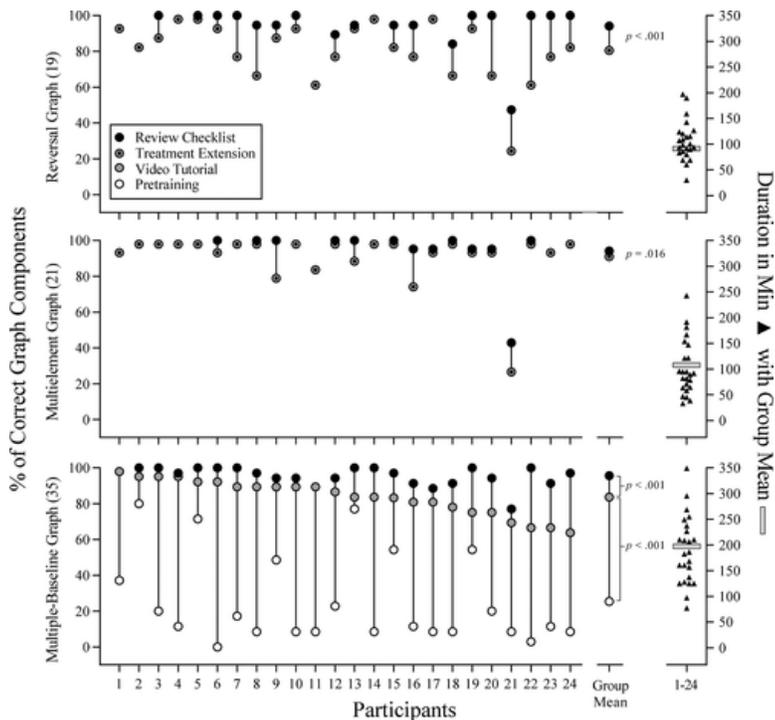
Note: Publication-quality features identified by Mitteer et al. (2018) are indicated with a superscript 1 and the automated features identified in this study are indicated with a superscript 2.

## Method

### Participants and Setting

Participants included 28 master's students in an applied behavior analysis program who were enrolled in the same 8-week course on the Measurement and Experimental Evaluation of Behavior. Students' graphs were scored for a grade because part of the course content involved teaching students to create single-case design graphs. We did not help students create graphs beyond giving them the resources described herein. Each student created graphs in any location they preferred using Excel on their personal MacOS or Windows computer that was equipped with a front-facing video camera. To be included in the study and to minimize technological difficulties, each student watched a 5-min technology test video<sup>1</sup> and (a) used Zoom<sup>2</sup> to record their computer screen and themselves using their front-facing video camera, (b) used the Google Chrome web browser to watch an unrelated video hosted on YouTube, and (c) opened Excel. Each student saved and uploaded their Zoom recording to an individualized folder on Microsoft OneDrive, a university-provided cloud-storage service. The technology test led to the exclusion of one student because Excel 2016 could not be

installed. One student was excluded due to unresolvable software errors in Excel that prevented graph creation, and two students were excluded due to not uploading Zoom recordings or Excel workbooks (i.e., the Excel database files) needed to score created graphs. In summary, data were included for 24 out of 28 students. Regarding graphing experience, 20 of the 24 students reported they had never used Excel to create the targeted graphs. Students 4 and 15 (see Figure 2) reported they had created one multiple-baseline graph, students 15 and 8 had created one and two multielement graphs, respectively, and students 15 and 23 had created one reversal graph.



**Figure 2** Percentage of Correct Graph Components and Duration to Graph Completion Across Graph Types and Conditions

*Note.* The primary y-axis parenthetical shows the number of components for each graph. The secondary y-axis depicts each student's duration to complete a graph using closed triangles and the students' mean duration using a grey bar. Group mean depicts  $p$ -values for significance tests to compare the pretraining and video-tutorial graphs (bottom panel, bottom value), video-tutorial and revised video-tutorial graphs (bottom panel, top value), multielement and revised multielement graphs (middle panel), and reversal and revised reversal graphs (top panel).

## Experimental Design

We used a quasi-experimental pretest–posttest design to systematically replicate Mitteer et al. (2018), who demonstrated functional control over the effects of their video tutorial using a multiple-baseline design across subjects. Our quasi-experimental design did not reveal or rule out the threat of testing, which would take the form of students improving their graphing accuracy due to additional attempts to create a graph in pretraining. We attempted to minimize the likelihood of undetected extraneous variables affecting our findings (Campbell & Stanley, 1963) by having students independently complete the pretraining and video-tutorial graphs in a 2-week period while recording themselves. Students' video recordings showed their computer screens and would have revealed if they used extra-experimental resources while creating the graph or had created the graph without watching the video tutorial, which would have suggested that extra-experimental history was a confound. Students' videos were reviewed using Auto-Movie-Thumbnailer<sup>3</sup> to view still video frames at 1-min intervals. No

students used extra-experimental resources, and all students watched the video tutorial while creating their graph.

## Materials

### Creating Target Graphs

See Supporting Information 1 for images of all target graphs. Three different multiple-baseline graphs were created that were identical to those created by Mitteer et al. (2018), but with title case for axis labels, slightly different maximum and major units on the y-axes, and different data point symbols. The y-axes' units and data point symbols were modified from the original graph due to software differences between Prism and Excel. Each graph contained two line-scatter plots and a column chart (i.e., a vertical bar chart) depicting dependent measures across baseline and intervention phases with three to 12 data points in each phase. Two graphs were used during pretraining (pretraining multiple-baseline graphs 1 and 2), and one graph was used during the video-tutorial training (video-tutorial multiple-baseline graph).

A treatment-extension multielement graph and a treatment-extension reversal graph were created with the publication-quality features identified by Mitteer et al. (2018) to test graphing accuracy to create novel graph types without a graph-specific video tutorial. The treatment-extension graphs entailed creating most of the same parts (e.g., phase lines) as the video-tutorial multiple-baseline graph, but in a different sequence; that is, the same parts of a graph were present across all graph types, but the sequence and number of parts differed across the multiple-baseline, multielement, and reversal graphs. The treatment-extension multielement graph took 196 steps to create using Excel and depicted five line-scatter plots described as demand, alone, attention, tangible, and play, with three to five data points per condition. The treatment-extension reversal graph took 260 steps to create using Excel and depicted a single line-scatter plot denoting responses per minute of destructive behavior across phases described as noncontingent attention or differential reinforcement of mands, with three to six data points in each phase.

### Creating the Video-Tutorial Training

A video tutorial was created for the same multiple-baseline graph as in Mitteer et al. (2018) using a Microsoft Surface Book 2 running Microsoft Windows 10 with Techsmith Camtasia version 9 to record the screen and a Zoom H1 Handy Recorder to record verbal instructions. Due to user interface differences between the MacOS and Windows versions of Excel in the *Select Data* menu, we used an Apple MacBook Air with Camtasia version 3 and the H1 recorder to record the steps related to the menu and embedded the recording into the video tutorial. As a result, students watched the same video tutorial regardless of operating system.

The 53 min 39 s video tutorial demonstrated how to create a multiple-baseline graph with nine publication-quality features identified by Mitteer et al. (2018) and five automated features. The automated features included (a) phase lines in panels that automatically adjust when the x-axis range is changed, (b) staggered phase lines between graph panels that perfectly align with phase lines in panels and automatically adjust when the x-axis range is changed, (c) text labels that automatically adjust when the x or y-axis ranges are changed, (d) data point labels and arrows that can be arbitrarily placed and automatically adjust when axes ranges are changed, and (e) a method to hide the y-axis line after floating the y-axis without using opaque shapes or text boxes (see Figure 1). The video tutorial contained 448 steps for Windows and 446 steps for MacOS users, was rendered at a resolution of 3,000 by 2,000 pixels, and was uploaded to YouTube.<sup>4</sup> Clickable hyperlinks with time stamps were embedded on the YouTube page below the video tutorial so students could easily navigate to different parts of the video tutorial, as done by Mitteer et al.

## Response Measurement and Interobserver Agreement

Scoring checklists were created to measure accuracy for the video-tutorial multiple-baseline graph and both treatment-extension graphs. The scoring checklist for the multiple-baseline graph contained 35 components and was adapted from the checklist used by Mitteer et al. (2018) with item changes specific to Excel (see Table 1). The scoring checklists for the multielement and reversal graphs contained 21 and 19 components, respectively (see Tables 2 and 3). The percentage of correct graph components was scored by visually inspecting a graph in the Excel workbook on the same operating system the student used to create the graph, summing the correct graph components, dividing by the total number of components, and multiplying the quotient by 100 to calculate a percentage. The duration in minutes to graph completion was scored by opening each video file using VLC media player and noting the total video time.

**Table 1.** Scoring Checklist for the Video-Tutorial Multiple-Baseline Graph

Number	Component Description
1	Top panel data values match target graph values
2	Middle panel data values match target graph values
3	Bottom panel data values match target graph values
4	“Self-Injurious Behavior” markers, symbols, and lines match target graph in all panels (except for line thickness)
5	“Pica” markers, symbols, and lines match target graph in all panels (except for line thickness)
6	Bar-graph is grey in all panels
7	Bar-graph width is slightly wider than data points in all panels
8	Proportions of all graph panels are approximately twice as long as they are tall
9	First-panel label matches target graph label (except for font size)
10	Second-panel label matches target graph label (except for font size)
11	Third-panel label matches target graph label (except for font size)
12	Labels are underlined in all panels
13	Arrangement of all panels matches target graph
14	X-axis range and tick marks match target graph and the x-axis is not obscured in any way
15	Labels for x-axis are present only on bottom panel
16	Zero values on the y-axis float slightly above the x-axis
17	X-axis label matches target graph (except for font size or presence of a text border) and is horizontally centered at the bottom of the graph
18	Primary y-axis range and tick intervals match target graph
19	Primary y-axis label matches target graph (except for font size or presence of a text border) and is vertically centered to the left of the graph
20	Secondary y-axis range and tick intervals match target graph
21	Secondary y-axis label matches target graph (except for font size or presence of a text border) and is vertically centered to the right of the graph
22	Baseline label matches target graph and is centered in the middle of the phase
23	Treatment label matches target graph and is centered in the middle of the phase
24	First-panel phase line is vertical, straight, solid, black, and is not crossed by a data path
25	Second-panel phase line is vertical, straight, solid, black, touches the first-panel phase line, and is not crossed by a data path
26	Third-panel phase line is vertical, straight, solid, black, touches the second-panel phase line, and is not crossed by a data path
27	Position of legend overlaps by approximately 50% or more when overlaid with the legend in the target graph
28	Legend is a rectangle with white background and black border
29	“Self-Injurious Behavior” legend text and symbol matches target graph

30	"Pica" legend text and symbol matches target graph
31	Bar-graph legend text and symbol matches target graph
32	Arrow points to the correct data point and does not touch the data point
33	Arrow label matches target graph
34	All font sizes match +2 sizes with the target graph (axis numbers and legend 7-11; axis labels 13-17; all other labels 9-13)
35	All line widths are less than 1.5 point

**Table 2.** Scoring Checklist for the Treatment-Extension Multielement Graph

Number	Component Description
1	Panel data values match target graph values
2	In Multielement "Play" and in Reversal "NCR and DRA" markers, symbols, and lines match target graph (except for line thickness)
3	"Demand" markers, symbols, and lines match target graph (except for line thickness)
4	"Alone" markers, symbols, and lines match target graph (except for line thickness)
5	"Attention" markers, symbols, and lines match target graph (except for line thickness)
6	"Tangible" markers, symbols, and lines match target graph (except for line thickness)
7	Proportions of all graph panels are approximately twice as long as they are tall
8	Zero values on the y-axis float slightly above the x-axis
9	X-axis range and tick marks match target graph and the x-axis is not obscured in any way
10	X-axis label matches target graph (except for font size or presence of a text border) and is horizontally centered at the bottom of the graph
11	Y-axis range and tick intervals match target graph
12	Y-axis label matches target graph (except for font size or presence of a text border) and is vertically centered to the left of the graph
13	Position of legend overlaps by approximately 50% or more when overlaid with the legend on the target graph
14	Legend is a rectangle with white background and black border
15	"Play" legend text and symbol matches target graph
16	"Demand" legend text and symbol matches target graph
17	"Alone" legend text and symbol matches target graph
18	"Attention" legend text and symbol matches target graph
19	"Tangible" legend text and symbol matches target graph
20	All font sizes match +2 sizes with the target graph (axis numbers and legend 7-11; axis labels 13-17; all other labels 9-13)
21	All line widths are less than 1.5 point

**Table 3.** Scoring Checklist for the Treatment-Extension Reversal Graph

Number	Component Description
1	Panel data values match target graph values
2	In Multielement "Play" and in Reversal "NCR and DRA" markers, symbols, and lines match target graph (except for line thickness)
3	Proportions of all graph panels are approximately twice as long as they are tall
4	Zero values on the y-axis float slightly above the x-axis
5	X-axis range and tick marks match target graph and the x-axis is not obscured in any way
6	X-axis label matches target graph (except for font size or presence of a text border) and is horizontally centered at the bottom of the graph

7	Y-axis range and tick intervals match target graph
8	Y-axis label matches target graph (except for font size or presence of a text border) and is vertically centered to the left of the graph
9	First label matches target graph and is centered in the middle of the phase
10	Second label matches target graph and is centered in the middle of the phase
11	Third label matches target graph and is centered in the middle of the phase
12	Fourth label matches target graph and is centered in the middle of the phase
13	First phase line is vertical, straight, solid, black, and is not crossed by a data path
14	Second phase line is vertical, straight, solid, black, and is not crossed by a data path
15	Third phase line is vertical, straight, solid, black, and is not crossed by a data path
16	Arrow points to the correct data point and does not touch the data point
17	Arrow label matches target graph
18	All font sizes match $\pm 2$ sizes with the target graph (axis numbers and legend 7-11; axis labels 13-17; all other labels 9-13)
19	All line widths are less than 1.5 point

The percentage of correct graph components for all graph types and phases was scored by a second observer for 34% of students. For correct graph components, an agreement was scored when observers scored the same graph component as correct or incorrect. Agreement was calculated for each graph by summing the total number of agreements, dividing by the total number of components, and multiplying the quotient by 100 to calculate a percentage. The mean and range of agreement for correct graph components in pretraining was 82% (range, 71%-100%), after video-tutorial training was 89% (range, 71%-97%), on the treatment-extension multielement graph was 89% (range, 67%-100%), on the treatment-extension reversal graph was 90% (range, 74%-100%), on the revised multiple-baseline graph was 98% (range, 94%-100%), on the revised treatment-extension multielement graph was 89% (range, 67%-95%), and on the revised treatment-extension reversal graph was 92% (range, 68%-100%). The duration in minutes to graph completion was independently scored by a second observer for 17% of students. An agreement was scored when observers scored the exact value for hours, minutes, and seconds for a given graph video. IOA was calculated for each graph by summing the total number of agreements, dividing by the total number of graphs scored, and multiplying the quotient by 100 to calculate a percentage. The IOA for duration to graph completion was 100%.

## Procedures

### General Procedures

All aspects of the project were conducted remotely using Zoom. At the start of each session, the student downloaded written instructions (see Supporting Information 2 for detailed instructions for each graph) and an Excel workbook for the assigned graph using the Canvas learning platform. Next, the student started a Zoom recording to capture their computer desktop and front-facing camera and opened the workbook. At the end of each session, the student saved the Excel workbook and Zoom recording and uploaded both files to OneDrive. Each student was asked to create graphs in a continuous manner, to watch the video tutorial with headphones, and to avoid seeking assistance. Students were given 1 week to create each graph and completed all review checklist graphs in a single week.

### **Pretraining**

Each student was randomly assigned to create either pretraining multiple-baseline graph 1 or pretraining multiple-baseline graph 2. The Excel workbook contained an image of the target graph and pre-entered data. In the written instructions, the student was instructed to end the session when they were satisfied with the graph or could no longer make progress in creating the graph.

### ***Video-Tutorial Training***

Each student watched the video tutorial and created the video-tutorial multiple-baseline graph, which was different than the pretraining graphs. The Excel workbook contained an image of the target graph and pre-entered data, a list of YouTube keyboard shortcuts, a list of the publication-quality features identified by Mitteer et al. (2018), and a list of the automated features. In the written instructions, we described how to change the video tutorial resolution to 1080p, how to use the hyperlinks below the video tutorial to navigate directly to a specific section of the video tutorial, and instructed the student to end the session when they were satisfied with the graph or could no longer make progress in creating the graph. At the end of the video tutorial, the experimenter showed how to make the target graph transparent and overlay it on the created graph to identify discrepancies and reviewed the list of publication-quality and automated features.

### ***Treatment Extension***

Procedures were identical with pretraining, except the student had access to the video tutorial for the multiple-baseline graph. The written instructions emphasized that although the video tutorial did not describe how to create a multielement or reversal graph, the student should use the hyperlinks below the video tutorial to navigate to relevant sections and use the same steps to create the multielement and reversal graphs. Supplemental instructions for the multielement graph were provided to change Excel's default setting for hidden and empty cells. The order in which each student created the graphs was randomly determined to reveal possible sequence effects related to practice. Half of the students created the multielement graph first and the reversal graph second, whereas the other half completed them in the opposite order.

### ***Review Checklist***

During this phase, students could use a review checklist to correct errors on the video-tutorial multiple-baseline, treatment-extension multielement, and treatment-extension reversal graphs. Some students were satisfied with their accuracy during the preceding phases and chose not to use the review checklist to correct errors. Review checklists were based on the corresponding scoring checklist for each graph type, but next to each checklist component there was a hyperlink to the relevant section of the video tutorial (see Supporting Information 3 for review checklists for each graph). Each student downloaded their Excel workbook containing a previously created graph, compared it to the target graph, and scored whether each component was correct by checking the corresponding box on the review checklist.

### ***Social Validity***

We developed a computer-based social-validity questionnaire composed of 18 questions using REDCap<sup>5</sup> and asked each student to complete it at the conclusion of all study phases (see Table 4). We asked eight objective questions about graphing experience, such as "In your best approximation, how many years of experience do you have using Excel to create graphs?" Using a 7-point Likert scale, we asked students to rate their satisfaction with the video tutorial and the review checklist and the likelihood of them recommending the video tutorial to other students and professionals, as asked by Tyner and Fienup (2015). For each Likert question, we also asked students to provide a rationale for their rating. On two items, we requested open-ended feedback about disliked aspects of and recommendations to improve the video tutorial and scoring checklist. We also asked students if they would have preferred to receive the review checklist immediately after using the video tutorial. We did not psychometrically validate our social-validity questions, but we incorporated recommendations by Fawcett (1991) to ask unbiased questions about the acceptability of the procedures and outcomes with the intent to use students' responses to improve the efficacy of the intervention.



5. Describe aspects of the review checklist you did not like or aspects you would change to improve it.								
Hyperlinks did not navigate to the precise part of the video tutorial to fix some errors.								
Even with the checklist and video tutorial, some errors could not be fixed.								
Treatment-extension checklist values differed from values shown in the video tutorial.								
6. Would you have preferred to use the review checklist immediately after the video tutorial?								
YES 88% (21)   NO 12% (3)								
7. How many years of experience do you have using Microsoft Excel to create graphs?								
0-1 years (14)   2-3 years (6)   4+ years (4)   Mean experience 1.5 years								
8. How many years of experience do you have creating graphs using software other than Microsoft Excel?								
0-1 years (19)   2-3 years (5)   Mean experience 0.6 years								

- *Note.* Data indicate the percentage of respondents ranking each item, with frequency of respondents in parentheses.

## Results

Individual student and group accuracy and duration to create graphs are depicted in Figure 2. Prism was used to conduct *t*-tests and to calculate effect sizes and confidence intervals. In pretraining, students' mean accuracy was 25.4% ( $n = 17$  scored between 0%-23% and  $n = 7$  scored between 37%-80%) and mean duration to create the graph was 63 min. A two-tailed, independent *t*-test used to compare mean accuracy between pretraining graph 1 ( $\bar{X} = 22.1\%$ ) and pretraining graph 2 ( $\bar{X} = 28.6\%$ ) revealed no significant difference ( $p = .539$ ), which suggests that neither pretraining graph was more difficult to create than the other. After watching the video tutorial and creating the video-tutorial multiple-baseline graph, students' mean accuracy increased to 85.7% ( $n = 10$  scored between 66%-85% and  $n = 14$  scored between 86%-100%), and mean duration to create the graph was 130 min. A one-tailed, paired *t*-test used to compare accuracy on the video-tutorial multiple-baseline graph versus the pretraining graph (mean of differences = 60.3%; 95% CI [50.3%, 70.4%])<sup>6</sup> revealed that students' mean accuracy after watching the video tutorial was significantly higher ( $p < .001$ ) with a large effect of the video tutorial ( $\eta^2_p = .87$ ). The effect size measure, partial eta squared ( $\eta^2_p$ ; also referred to as  $R^2$ ), is the percentage of explained variation between two group means<sup>7</sup> (GraphPad, 2014). Next, 22 of 24 students used the review checklist and video tutorial to correct errors, resulting in an increase in their mean accuracy to 95.7% ( $n = 2$  scored a 77% and 89%, and  $n = 20$  scored above 91%). A one-tailed, paired *t*-test used to compare accuracy between the review checklist with video tutorial versus the video tutorial alone (mean of differences = 10.9%; 95% CI [6.9%, 15.0%]) revealed that students' accuracy after using the review checklist was significantly higher ( $p < .001$ ) with a large effect of the review checklist ( $\eta^2_p = .60$ ). The mean duration to watch the video tutorial, create the graph, and use the review checklist to correct errors was 190 min ( $n = 13$  took between 77-186 min and  $n = 11$  took between 207-349 min). In summary, all students created a multiple-baseline graph more accurately using the video tutorial, and students' mean accuracy further improved after using the review checklist.

On the treatment-extension multielement graph, students' mean accuracy using the video tutorial was 93.1% ( $n = 5$  scored between 29%-90% and  $n = 17$  scored between 95%-100%), using the review checklist with the video tutorial was 94.1% ( $n = 1$  scored 43% and  $n = 12$  scored between 95%-100%), and they created the graphs in a mean duration of 102 min. Note that 11 of 24 students chose not to use the review checklist because they were satisfied with their accuracy obtained with only the video tutorial. A one-tailed, paired *t*-test used to compare accuracy between the review checklist with the video tutorial versus the video tutorial alone (mean of differences = 5.1%; 95% CI [0.5%, 9.6%]) revealed that students' ( $n = 13$ ) accuracy was significantly higher ( $p = .016$ ) with the checklist than watching the video tutorial alone with a medium effect of the review checklist ( $\eta^2_p = .33$ ).

On the treatment-extension reversal graph, students' mean accuracy after watching the video tutorial was 82.5% ( $n = 10$  scored 26%-79% and  $n = 14$  scored between 84%-100%), using the review checklist with the video tutorial was 94.2% ( $n = 3$  scored 47%-89% and  $n = 15$  scored between 95%-100%), and they created the graphs in a mean duration of 106 min. A one-tailed, paired *t*-test used to compare accuracy between the review checklist with the video tutorial versus the video tutorial alone (mean of differences = 14.3%; 95% CI [9.1%, 19.6%]) revealed that students' ( $n = 18$ ) accuracy was significantly higher with the review checklist ( $p < .001$ ) than watching the video tutorial alone with a large effect of the review checklist ( $\eta^2_p = .66$ ). A two-tailed, independent *t*-test used to reveal possible order effects during the treatment extension revealed no significant differences in accuracy between students who completed the multielement graph first ( $p = .950$ ) versus those who completed the reversal graph first ( $p = .168$ ). Therefore, students' higher mean accuracy on the multielement graph is likely due to the lower number of steps (196) relative to the reversal graph (260 steps).

An error analysis of the video-tutorial multiple-baseline graph revealed that 58% of students made errors connecting, aligning, and formatting phase lines. More than half of students (54%) made errors manually aligning or formatting a grey glyph on the legend to depict the symbol for the column charts, which can only happen when creating a graph that includes a line-scatter plot and a column chart. More than half of students (54%) made errors related to the x- and y-axes, including not removing unneeded tick mark labels, not removing a part of the y-axis that extends below the x-axis, and mistakenly hiding parts of the x-axis. Nearly half of students (46%) misaligned or made spelling errors on text labels for the axes. An error analysis of the reversal graph revealed that 67% of students misaligned the text labels for phases. An error analysis of the multielement graph was not conducted because most students made three or fewer errors.

Students' social-validity ratings are depicted in Table 4. Most students expressed satisfaction with the video tutorial ( $M = 5.0$ ; range, 2-7) and would recommend it to individuals learning to graph ( $M = 4.9$ ; range, 1-6). More students expressed satisfaction with the review checklist ( $M = 5.4$ ; range, 2-7), and they overwhelmingly (88%) indicated they would have preferred to use the review checklist to review the accuracy of graphs immediately after creating them. Students' mean experience in creating graphs using Excel was 3.5 years for the four students who had previously created multiple-baseline, multielement, or reversal graphs, and 1.2 years for the remaining students. Students' mean experience in creating graphs using other graphing software (e.g., Prism) was 0.6 years.

## Discussion

The video tutorial for Excel produced large and statistically significant increases in graphing accuracy for 24 students who created a multiple-baseline graph with publication-quality features. This outcome systematically replicates Mitteer et al. (2018), whose participants scored 86%-100% while using a video tutorial for Prism. As a comparison, students in our study scored between 66%-100%, indicating that the Excel video tutorial produced less consistent effects than the Prism video tutorial. However, our sample size was six-fold larger than Mitteer et al., and it is unknown whether similar intersubject variability would have been observed with a larger sample in Mitteer et al. Nevertheless, after students corrected errors on their multiple-baseline graph using a review checklist, their mean accuracy increased from 86% to 96%, with 21 of 22 students scoring between 86%-100%. Students also took longer to create the graph ( $M = 130$  min) than participants in Mitteer et al. ( $M = 45$  min). We then extended Mitteer et al. by testing the generality of the video tutorial on students' accuracy in creating multielement and reversal graphs, and their respective mean accuracy was 93% and 82%. After students fixed errors using the review checklists, mean accuracy improved to 94% on both graphs. Students reported moderate to high satisfaction with the video tutorial and review checklists, and they overwhelmingly reported a preference to use the checklists immediately after creating each graph. We did not control where and when students created graphs, and students completed the study without meeting in-person with the first author. The sole methods of communication and file sharing involved one Zoom video conference, multiple emails, the Canvas learning platform, and OneDrive, which provides support for the generality and external validity of our procedures and indicates they may be suitable for in-person and distance learning courses.

The greater variability in graphing accuracy and the increased duration to create the multiple-baseline graph in our study compared to Mitteer et al. (2018) was likely due to the higher number of steps (446) and longer viewing duration (54 min) of the Excel video tutorial compared to the steps (226) and viewing duration (37 min) of the Prism video tutorial. These differences were due to how Excel and Prism inherently operate and our decision to include additional steps to create automated phase and staircase lines, text labels, and arrows. The higher number of steps in the Excel video tutorial provided more opportunities to make errors, and we asked students to create each graph in a single sitting to minimize threats to internal validity. Both factors may have resulted in increased fatigue or inattention for at least one student who appeared to fall asleep several times

while creating the graph. To address this, individuals should be allowed to create graphs across multiple sittings. Regarding the longer viewing duration, researchers could compare performance and social validity across automated versus manual methods to insert phase lines, text labels, and arrows.

The error analysis of the multiple-baseline graph and students' preferences informed modifications to the review checklist and when it should be used. We modified review checklist items to recommend individuals (a) check for errors using Excel's zoom function at 150% by closely inspecting the legend, axes, and phase lines. We added a review checklist item to (b) use Excel's spell check function to identify spelling errors in text labels because Excel does not automatically underline spelling errors. Both (a) and (b) modifications are supported by our finding that the review checklist helped students improve their accuracy, likely because it guided them in visually inspecting the graph. Considering this finding and students' preferences, we recommend (c) individuals access the review checklist concurrently with the video tutorial and use it during graph creation or immediately afterward; research evaluating whether this recommendation could further increase graphing accuracy is warranted. Finally, we also suggest it would be helpful if individuals use a second monitor or connected device (e.g., large tablet, television) to watch the video tutorial and create the graph at the same time. Researchers could evaluate the performance and social validity of individuals creating graphs while using a second monitor to watch the video tutorial. We did not empirically evaluate (a), (b), or (c) in the present study.

Although using the review checklist helped 21 of 22 students increase their accuracy by a mean of 10 percentage points, a few errors persisted for most students. In response, we recommend (a) an automated method to create the column chart to eliminate errors caused by manually inserting the legend glyph and (b) using an opaque shape to manually hide unneeded tick mark labels and y-axes parts to eliminate errors caused by failing to hide these unneeded features or mistakenly hiding other parts of the graph. Regarding (a), recall that we taught students to create the column chart using an error bar and to manually insert a legend glyph because Excel cannot properly combine a line-scatter plot and column chart in the same graph, as needed to replicate the graphs by Mitter et al. (2018). For example, if two line-scatter plots are graphed and one line-scatter plot is changed to a column chart, Excel will properly display an x-axis for the column chart, but not for the line-scatter plot. In addition, column charts cannot have a floating y-axis. To eliminate the need to manually insert the legend glyph, we identified an alternative method to create the column chart as an automated feature on a line-scatter plot, which allows Excel to automatically create the legend. Regarding (b), we taught students an automated method to hide unneeded tick-mark labels and y-axes parts (Deochand et al., 2015), but despite the accurate completion of the corresponding steps in the video tutorial, inadvertently resizing the graph panels or the chart area inside a panel can cause parts of the graph to be hidden. For an alternative to the automated method, an opaque shape can be used to hide unneeded parts of the graph using fewer steps than the automated method (Chok, 2018; Lo & Konrad, 2007; Lo & Starling, 2009). We created video tutorials for both alternative methods and placed links to them on the YouTube page below the video tutorial. We did not empirically evaluate these alternative methods in the present study, so researchers should aim to replicate our obtained results using the alternative video tutorials.

Adopting several of the students' suggestions may enhance the social validity of the video tutorial and the context in which training is delivered. Most students (19 out of 24) used MacOS, and eight students disliked the Windows-based video tutorial due to cosmetic differences in the user interface (e.g., the menu to reselect data looked notably different). Satisfaction may improve if individuals could access a video tutorial depicting the MacOS version of Excel, and researchers could evaluate performance and social validity differences between groups of students using operating system-specific video tutorials. Four students requested video tutorials specific to the multielement and reversal graphs, but we did not provide these because we evaluated the generality of the multiple-baseline video tutorial to create multielement and reversal graphs. Researchers should evaluate performance and social validity differences between the procedures used in the present study

versus individuals creating each graph using a specific video tutorial. On the review checklists, two students disliked that some hyperlinks with time stamps linked to a general section of the video tutorial instead of the precise steps needed to fix parts of a graph, so we changed all hyperlinks and time stamps to link to the precise steps. Six students wanted additional graphing resources, help from the experimenter, or in-person instruction with the video tutorial. We chose not to provide additional assistance to our students to replicate the methods by Mitteer et al. (2018) and to reduce threats to internal validity, but in an academic or training setting it seems prudent to give individuals choices for how instructions are presented. For example, some individuals may prefer to use the video tutorial and review checklists alone and at a place and time most convenient to them. Other individuals may prefer to use the video tutorial and review checklists in-person with an instructor who can provide technical support (Deochand et al., 2015) or assistance with difficult steps. If an instructor cannot provide in-person assistance, Zoom can be used for video conferencing and has screen sharing and remote-control capabilities. For example, an instructor could schedule a Zoom meeting with an individual to view their graph in real time, observe their graphing behaviors (i.e., what buttons they are clicking or values they are inputting), or even assume control of their computer to model correct graphing behaviors. Finally, although we did not test the video tutorial with Excel 2013 and lower, we confirmed it could be used with the Windows and MacOS versions of Excel 2019 and Excel Office 365, and we will update the video tutorial for practitioners, students, and teachers if future Excel updates impact its functionality.

Our use of a quasi-experimental pretest–posttest design with a single pretraining measure did not rule out the threat of repeated testing, but testing effects seem unlikely to explain students' large increases in accuracy. For example, in the study conducted by Mitteer et al. (2018), no participants reached mastery criteria until they viewed the video tutorial, even after up to five repeated baseline measures. To better control for threats of history and testing, researchers who systematically replicate this study should consider adopting the design used by Fisher et al. (2003; Study 3), which resembled a multiple probe design across subjects using groups of individuals. Individuals would be randomly assigned to Group A and Group B. During the first session, both groups would attempt to create graph 1. During the second session, one group would be randomly assigned (e.g., Group A) to use the video tutorial to create graph 2 while Group B would attempt to create graph 1 for a second time. During the third session, Group B would use the video tutorial to create graph 2 while Group A used the video tutorial to create graph 2 for a second time. This experimental design may be amenable to the practical constraints of embedding research in an academic course.

<sup>1</sup> <https://youtu.be/XcWSIH0jAuk>

<sup>2</sup> Zoom (<https://zoom.us>) was approved by the University of Nebraska Medical Center for video conferences involving protected health information. We used Zoom because it can record the screen and front-facing camera to measure and evaluate graphing behaviors and because individuals can use it for free.

<sup>3</sup> <http://www.amt.cc>

<sup>4</sup> To view the video tutorial that students used for this evaluation, please navigate to <https://youtu.be/lmcsjRR2P4> Practitioners, students, and teachers may prefer this updated version of the video tutorial with edits to increase usability: <https://youtu.be/PD-s4MS8ZPw>

<sup>5</sup> Study data were collected and managed using the REDCap electronic data capture tools hosted at the University of Nebraska Medical Center. Service and support is provided by the Research Information Technology Office, which is funded by the Vice Chancellor for Research.

<sup>6</sup> The 95% confidence interval (95% CI) with lower and upper confidence limits shown in brackets, indicates 95% confidence that the population mean lies within the confidence limits (Motulsky, 2017). For example, if 95% CI [50%, 70%], then there is 95% confidence that the population mean is between 50% - 70%.

<sup>7</sup> For example, if  $\eta^2_p = .20$ , then 20% of the variation between group means can be attributed to a difference between groups, such as the difference that results from the intervention's effects.

## References

- Arsenault, D. J., Smith, L. D., & Beauchamp, E. A. (2006). Visual inscriptions in the scientific hierarchy: Mapping the "treasures of science." *Science Communication*, **27**(3), 376– 428. <https://doi.org/10.1177/1075547005285030>
- Behavior Analyst Certification Board (2012). BCBA® and BCaBA® Fourth Edition Task List. Retrieved June 2, 2019, from <https://www.bacb.com/wp-content/uploads/2017/09/160101-BCBA-BCaBA-task-list-fourth-edition-english.pdf>
- Behavior Analyst Certification Board (2018). RBT™ Second Edition Task List. Retrieved June 2, 2019, from [https://www.bacb.com/wp-content/uploads/RBT-2nd-Edition-Task-List\\_181214.pdf](https://www.bacb.com/wp-content/uploads/RBT-2nd-Edition-Task-List_181214.pdf)
- Bitglass (2018). Bitglass Cloud Adoption 2018. Retrieved from [https://pages.bitglass.com/FY18BR-CloudAdoption\\_LP.html](https://pages.bitglass.com/FY18BR-CloudAdoption_LP.html)
- Campbell, D. T., & Stanley, J. C. (1963). *Experimental and quasi-experimental designs for research*. Houghton Mifflin.
- Chok, J. T. (2018). Creating functional analysis graphs using Microsoft Excel® 2016 for PCs. *Behavior Analysis in Practice*, **12**(1), 265– 292. <https://doi.org/10.1007/s40617-018-0258-4>
- Deochand, N. (2017). Automating phase change lines and their labels using Microsoft Excel®. *Behavior Analysis in Practice*, **10**(3), 279– 284. <https://doi.org/10.1007/s40617-016-0169-1>
- Deochand, N., Costello, M. S., & Fuqua, R. W. (2015). Phase-change lines, scale breaks, and trend lines using Excel 2013. *Journal of Applied Behavior Analysis*, **48**(2), 478– 493. <https://doi.org/10.1002/jaba.198>
- Dubuque, E. M. (2015). Inserting phase change lines into Microsoft Excel® graphs. *Behavior Analysis in Practice*, **8**(2), 207– 211. <https://doi.org/10.1007/s40617-015-0078-8>
- Fawcett, S. B. (1991). Social validity: A note on methodology. *Journal of Applied Behavior Analysis*, **24**(2), 235– 239. <https://doi.org/10.1901/jaba.1991.24-235>
- Fisher, W. W., Kelley, M. E., & Lomas, J. E. (2003). Visual aids and structured criteria for improving visual inspection and interpretation of single-case designs. *Journal of Applied Behavior Analysis*, **36**(3), 387– 406. <https://doi.org/10.1901/jaba.2003.36-387>
- Fuller, T. C., & Dubuque, E. M. (2019). Integrating phase change lines and labels into graphs in Microsoft Excel®. *Behavior Analysis in Practice*, **12**(1), 293– 299. <https://doi.org/10.1007/s40617-018-0248-6>
- GraphPad (2014, April 24). *R2 of unpaired and paired T test results*. <https://www.graphpad.com/support/faqid/918/>
- Kubina, R. M., Kostewicz, D. E., & Datchuk, S. M. (2008). An initial survey of fractional graph and table area in behavioral journals. *The Behavior Analyst*, **31**(1), 61– 66. <https://doi.org/10.1007/BF03392161>
- Lo, Y., & Konrad, M. (2007). A field-tested task analysis for creating single-subject graphs using Microsoft® Office Excel. *Journal of Behavioral Education*, **16**(2), 155– 189. <https://doi.org/10.1007/s10864-006-9011-0>
- Lo, Y., & Starling, A. L. P. (2009). Improving graduate students' graphing skills of multiple baseline designs with Microsoft® Excel 2007. *The Behavior Analyst Today*, **10**(1), 83– 121. <https://doi.org/10.1037/h0100666>
- Mitteer, D. R., Greer, B. D., Fisher, W. W., & Cohrs, V. L. (2018). Teaching behavior technicians to create publication-quality, single-case design graphs in Graphpad Prism 7. *Journal of Applied Behavior Analysis*, **51**(4), 998– 1010. <https://doi.org/10.1002/jaba.483>
- Motulsky, H. (2017). *Intuitive biostatistics: A nonmathematical guide to statistical thinking*. Oxford University Press.
- Tufte, E. (2001). *The visual display of quantitative information*. Graphics Press.
- Tyner, B. C., & Fienup, D. M. (2015). A comparison of video modeling, text-based instruction, and no instruction for creating multiple baseline graphs in Microsoft Excel. *Journal of Applied Behavior Analysis*, **48**(3), 701– 706. <https://doi.org/10.1002/jaba.223>
- Vanselow, N. R., & Bourret, J. C. (2012). Online interactive tutorials for creating graphs with Excel 2007 or 2010. *Behavior Analysis in Practice*, **5**(1), 40– 46. <https://doi.org/10.1007/BF03391816>