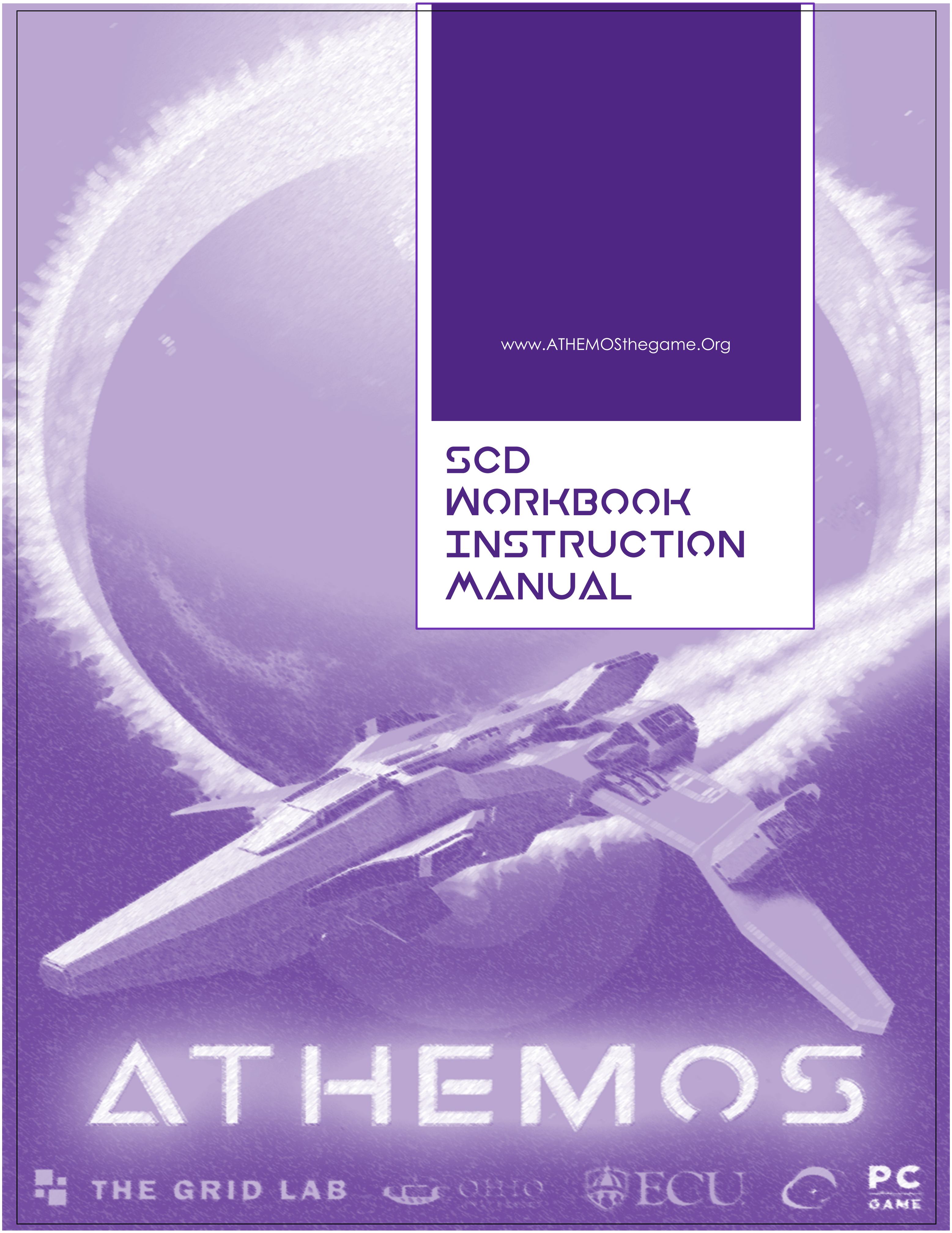
www.ATHEMOSthegame.Org

SCD WORKBOOK INSTRUCTION MANUAL



Disclaimer

The Excel workbook is intended for use when assessing rates of improvement or comparing changes in average performance over time during the ATHEMOS interventions. No workbook can substitute for sound judgment (or common sense), so readers are cautioned to consider their intervention, the target child, and their experiences before drawing any conclusions.

Also, the workbook will not work on every computer. This tool was created in Excel 2013 using Windows and might not work in another version of the program or on another operating system. If users cannot open the file, our suggestion is to update to Excel 2013 (or newer), and/or use the workbook on a computer running Windows. Even still, we cannot guarantee it will work for everyone.

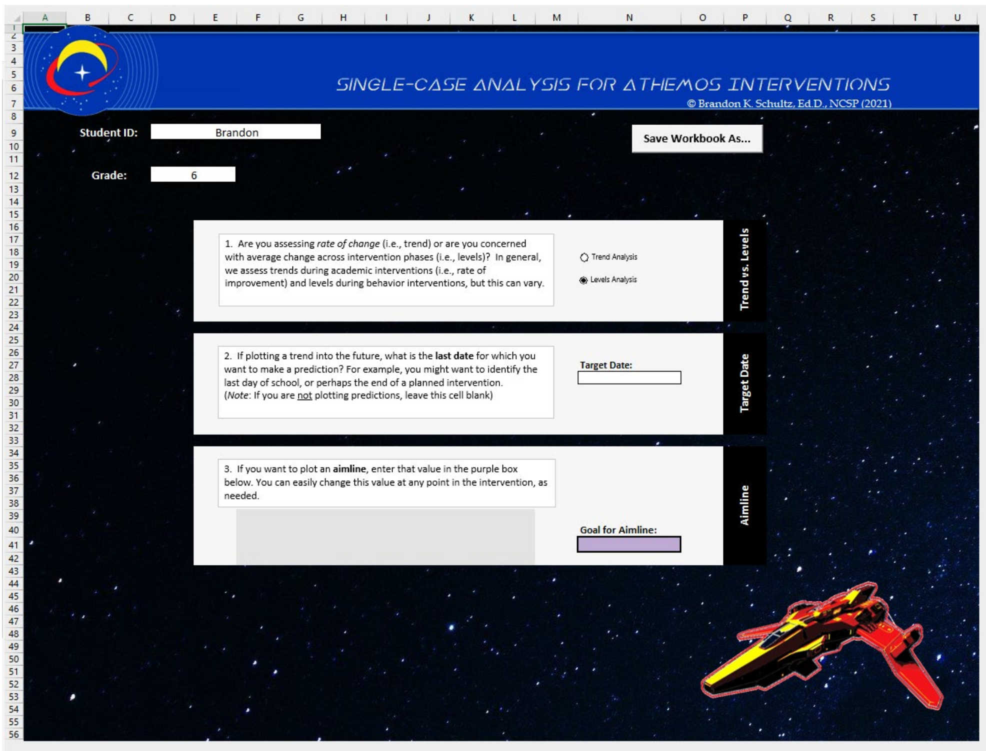
INDEX

GETTING STARTED

START TAB

When users open the workbook, they will see the "Start tab" (see Figure 1).

FIGURE 1. START TAB



The first time opening the workbook users are likely to receive a warning about "macros." The workbook contains code necessary for calculating effect sizes or supporting many of the features in the workbook. To our knowledge, there is no risk to using the workbook.

There are a few things to note. First, the normal Excel "ribbon" is hidden at the top of the screen and the worksheets are locked so that only a few cells will allow users to make changes. This is to protect the underlying elements. Note that the workbook is not password protected, so users will need to protect the confidentiality of student clients by keeping the file in a safe location and identifying clients without using full names or other protected information (e.g., addresses).

When starting a new intervention, we recommend the following steps:

1. Enter the child's code name in cell **D9**.
2. Enter the child's grade in cell **D12**.
3. Click the "**Save Workbook As…**" button, rename the file, and save (each child should have his or her own workbook).

Trends of Levels? Once the workbook is saved, users will need to make some decisions about how to use the workbook. The first decision is to determine whether to consider outcomes in terms of trends (i.e., rate of improvement) or in terms of levels. As explained on the screen, trends are generally considered when assessing growth over time and levels in behavioral interventions, but that can vary. By clicking either of the two options, the worksheet tabs that are visible at the bottom of the screen will change, but users can toggle between these two options to see the results of both the trend and level analyses after data are entered.

Target Date? Next, users will need to identify a target date in cell N28 to make predictions. Typically, this only applies to trend analyses. If there is no target date in mind, users can leave this cell blank, and the target will be the last date of data collection. Note that the format for dates is **mm/dd/yyyy**.

Goal? Finally, users will need to identify a goal in cell N41 to plot an aimline. The aimline option is only available in the trend analysis and is ignored in the level analysis.

ENTERING DATA

AB_DATA TAB

To enter data, click on the AB_Data tab at the bottom of the screen. The workbook can manage up to four phases in a single-case design, but we programmed it with the classic AB or ABAB designs in mind. For alternative designs (e.g., B-study, BAB design), users will need to interpret their results accordingly.

Note that data are entered in columns, with dates entered sequentially in column A and each phase starting in a new column, beginning with column C (see Figure 2). Also note that *copy* and *paste* functions are disallowed because these features can damage the underlying macros. Users will have to enter and delete data one cell at a time. Each row is intended to hold one day of data, so measures collected more than once per day will need to be averaged.

FIGURE 2. AB_DATA DATA ENTRY SCREEN

	A	C	D	E	F	G
1	Dates	Baseline 1	Intervention 1	Baseline 2	Intervention 2	
2	1/11/2016	42.5				Export AB Data into .txt
3	1/13/2016	48.0				
4	1/15/2016	38.0				
5	1/18/2016	43.0				Export All Data into .csv
6	1/20/2016	50.0				
7	1/25/2016	42.0				
8	1/27/2016	39.0				
9	1/29/2016	45.0				
10	2/3/2016		40.0			
11	2/5/2016		50.0			
12	2/10/2016		42.0			
13	2/12/2016		52.0			
14	2/22/2016		65.0			
15	2/24/2016		60.0			
16	2/29/2016		64.0			
17	3/2/2016		61.0			
18	3/4/2016		67.0			
19	3/9/2016		62.0			
20	3/11/2016		68.0			
21						
22						
23						

Start | AB_Data | TrendResults

In figure 2, the example data are configured for a simple AB design. Note that the baseline data are entered in column C. The intervention then starts on 2/3 and those data are entered in column D. If the interventionist returns to baseline, those data would be entered in Column E, and so forth. Be careful not to overlap phases because it can lead to inaccurate results.

The workbook is programmed to calculate regression equations and effect sizes regardless of the intervals between measurement occasions—just enter the dates as they occur (e.g., **1/11/16**, **1/13/16, 1/15/16…**). The only concern is that the dates are entered sequentially, from earliest to latest, and that the data to the right are accurately aligned with those dates.

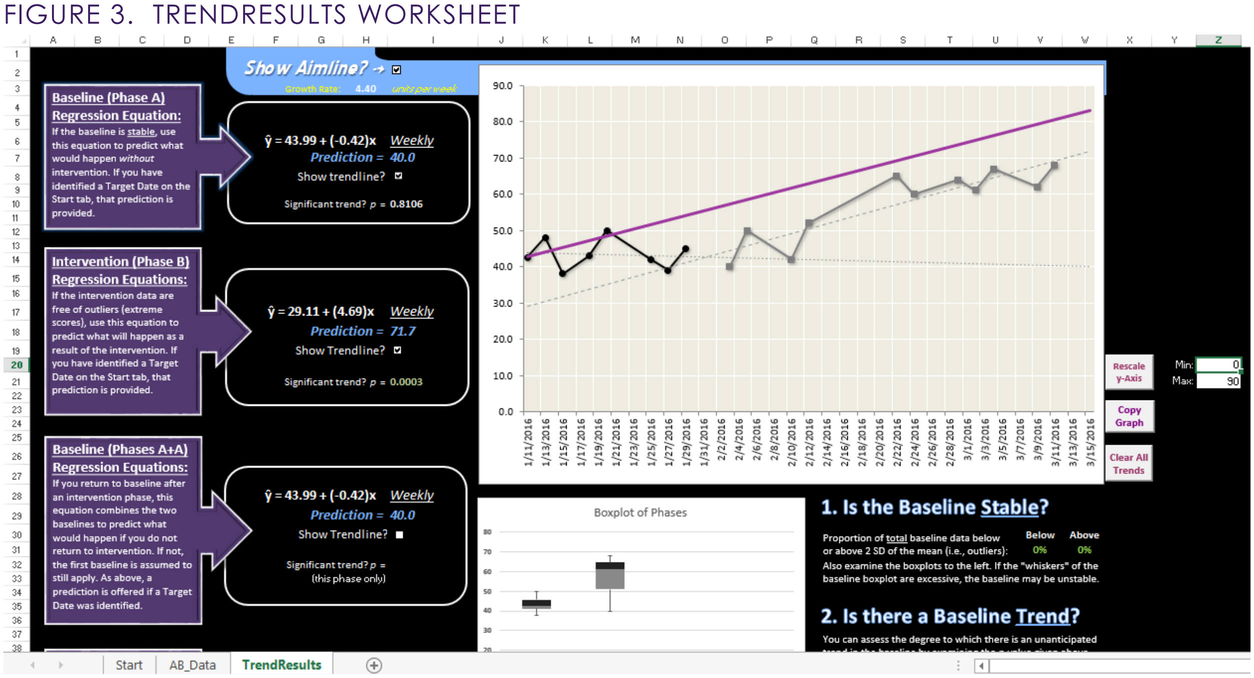
Also note that there are two buttons (top right) that allow users to export data from the workbook into either text files or comma-separated files. The former makes it easy to copy data into a stats package, like *R*,[i] and the latter makes it easy to move data into a separate Excel workbook to do additional analyses or create different graphs.

EVALUATING TRENDS

TRENDRESULTS TAB

If assessing *rate of improvement*, the next tab at the bottom of the screen is labelled "TrendResults." The TrendResults worksheet is shown in Figure 3.

FIGURE 3. TRENDRESULTS WORKSHEET



Trend Results. For each phase with data, OLS regressions equations are calculated in the rounded boxes. Note that beginning with the second baseline phase (third phase), the two baselines are combined to offer one regression equation. So, in the example shown in Figure 3, the AB design only has one baseline, so the same regression estimates are provided for the Phase A and what is termed Phases A+A. If the interventionist returned to baseline, the regression equation for Phases A+A would update and include both baselines in the calculation. The same assumptions are made when there are two intervention phases (see estimates for Phases B+B).

Progress toward Goals. To graphically see the trends, click the checkboxes labelled 'Show Trendline?" provided for each phase. Users can remove trendlines by unchecking these boxes or clicking the "Clear All Trends" button to the right of the graph. To see the aimline, click the checkbox at the top of the screen (near cell **I2**). The aimline is based on the goal identified on the **Start tab in cell N42**, and it can be removed by either unchecking the box or clicking the "Clear All Trends" button. Note that the graph will adjust to include the target date if one is provided in **cell N28 of the Start tab**. When using this option, the necessary growth rate per week is provided in cell H3. Note that in this example, the target date is set at 3/15/16.

INTERPRETING TREND RESULTS

It is beyond the scope of these instructions to go into every possible detail about trend analyses, so users should refer to books and articles on RTI decision-making. For our purposes here, however, it is important to point out what the workbook can and cannot do.

Stability of the Baseline. First, the TrendResults worksheet shows the proportion of data that are "outliers" (beyond 2SD from the mean) in cells **V32** and **W32**. If either of these estimates exceed 5% it suggests that the baseline may not be stable. If possible, continue to collect baseline data. If this is not possible, consider using nonparametric tests of the trend (explained below).

The spreadsheet also provides a boxplot of the phases. If any of the boxplots appear odd (e.g., long or nonexistent whiskers), users might consider nonparametric tests of the trend and ignore the regression-based results. Please note, however, that the outlier value and boxplots are rough estimates. If the results of the analysis do not comport with the interventionist's experience and the child, then consider using an alternative stats package (e.g., *R*) for additional tests.

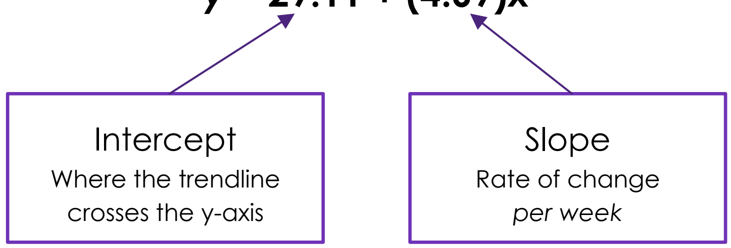
Unanticipated Trends in the Baseline. Second, the TrendResults worksheet can help to determine whether there are unanticipated trends in the baseline phase(s). Baseline data that are trending upward or downward prior to (or after) intervention can complicate the interpretation of the results because what appears to be an intervention effect may simply be an artifact of external developments (e.g., improvements already underway in the classroom). One test of this possibility is the *p*-value associated with the slope of the baseline regression line. These results are provided for each phase in the TrendResults worksheet, but the most important values are associated with the baseline phases—if either of these values are significant ($p < .05$), then users must proceed cautiously. Ideally, users will test their baseline data *before* starting intervention, which would allow them to collect additional data when there is a significant trend. (Note, however, that significant trends in the *undesired* direction—for example, a significant *decline* in oral reading fluency—would hasten the need to begin intervention.)

A second "test" of trend can be achieved by plotting the trendline for A or A+A. When the trendline is plotted, does the prediction appear unrealistic? In other words, does it predict unachievable values in the desired direction (e.g., oral reading fluency greater than 150)? If so, there may be a meaningful trend in the baseline, <u>even if the *p*-value > 0.05</u>. The best remedy in such cases is, again, to collect additional baseline data. If that is impossible, users might consider nonparametric tests of the intervention and ignore the regression-based results.

Autocorrelation. Third, the TrendResults worksheet tests for autocorrelation, which can distort the results. A bar graph shows the autocorrelation estimate for each phase, along with the upper and lower critical values identified as red lines. If any of the bars exceed the critical values (i.e., go above or below the red lines), significant autocorrelation has occurred. In those instances, users will probably want to use nonparametric tests of the trend. Even visual analysis of the data is suspect when autocorrelation is significant, so be careful! Interventionists make this mistake all the time!

Regression Equation. If the data appear to be stable, free of baseline trends, and not significantly autocorrelated, report the intervention results using the OLS regression estimates provided in the worksheet. The equations are interpreted as follows:

$$\hat{y} = 29.11 + (4.69)x$$



Note that the slope estimates produced by the workbook are *per week*. If users want to calculate the daily rate of improvement, divide this value by 7; if users want to calculate the monthly rate of improvement, multiply this value by 4. What is most important, of course, is that the slope estimate exceeds the growth rate needed to hit the goal, provided in cell **H3** (also per week). In the example depicted in Figure 4, the growth rate of 4.69 exceeds the needed growth rate of 4.40. In this instance, users would conclude that the child is responding to the intervention and that referral to intensive services appear unnecessary at this time. If, however, the slope does not exceed the needed growth rate (and assuming the intervention is implemented with fidelity), interventionists might change or add new services.

WHAT IF THERE ARE PROBLEMS?

The last section of the TrendResults worksheet provides the trend estimate using the Tukey method. The Tukey method is used when there are obvious problems with the OLS regression described above. The Tukey method estimates a trend based on median baseline performance and the median of the *last three data* of the intervention phase. It ignores all data in the middle of the trend, so it should only be used when the assumptions of OLS regression are violated.

When users click the "Show Tukey Trendline" checkbox, the Tukey trendline is plotted in the graph. Note that the intercept and slope estimates of the Tukey estimate are provided in cells **L70** and **L72**, respectively.

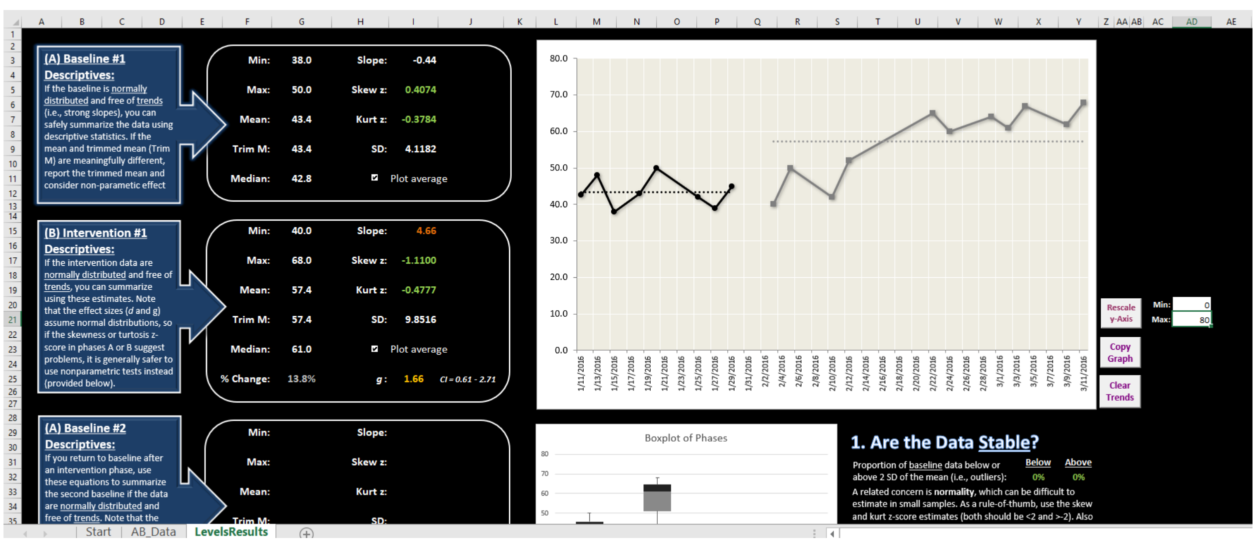
There are other alternatives when the assumptions of OLS regression have been violated, but those options are beyond the capabilities of Excel. If the Tukey estimate appears problematic, users can examine the nonparametric level comparisons (PND and PEM) in the LevelsResults worksheet, but these are generally poor options to assess the *rate* of improvement. When users' questions or data exceed what this workbook can do, the best option is to use a stats package designed to handle single-case data. (In those instances, we recommend *R* because it is free and offers current approaches to single-case analysis, but it does require some coding. Please visit SchoolPsychologyTech.Org for free code that might help in that situation.)

EVALUATING LEVELS

LEVELSRESULTS TAB

If assessing change in *average performance* across phases, the next tab at the bottom of the screen is labelled "LevelsResults." The LevelsResults worksheet is shown in Figure 5.

FIGURE 4. LEVELSRESULTS WORKSHEET



For each phase with data, descriptive statistics are calculated in the rounded boxes. Note that beginning with the first intervention phase, a "standardized mean difference" effect size is provided; namely, **Hedge's *g***. This estimate (1.66 in Figure 4) summarizes the difference between two phases, divided by the pooled standard deviation adjusted for differing numbers of observations in each phase (as is typical in single-case designs). Hedge's *g* is further corrected for small sampling bias, as recommended by Durlak (2009), because this is also true in most single-case designs (i.e., few measurement occasions). The full formula is as follows:

$$g = \frac{M_2 - M_1}{SD_{pooled}} \times \left(\frac{N-3}{N-2.25}\right) \times \sqrt{\frac{N-2}{N}}$$

Where M_2 is the mean of the second phase, M_1 is the mean of the first phase, and N is the total number of measurement occasions. The pooled standard deviation is calculated as:

$$SD_{pooled} = \sqrt{\frac{[(SD_2)^2(n_2-1)] + [(SD_1)^2(n_1-1)]}{(n_2+n_1)-2}}$$

Where n_1 and n_2 are the sample sizes of the first and second phases, respectively. Similarly, SD_1 and SD_2 are the standard deviations of the first and second phases, respectively.

Each effect size (*g*) on the LevelsResults worksheet compares the current phase with the immediately preceding phase. For example, the estimate on row 25 compares the first intervention phase with the first baseline phase. If the interventionist returned to baseline, the relevant effect size estimate on row 39 would compare the second baseline with the first intervention, and so forth. One total effect size is given on row 60 but note that it combines the A-phases and compares them to the combined B-phases (assuming an ABAB design), so it is not appropriate when an alternative design is used (e.g., ABC).

To graphically see the means of each phase, click the checkboxes labelled "Plot average" provided for each phase. Users can remove the means by unchecking these boxes or clicking the "Clear Trends" button to the right of the graph (users might have to click this button multiple times).

INTERPRETING LEVELS RESULTS

It is beyond the scope of this instruction booklet to go into every possible detail about level comparisons, so users should refer to books and articles on RTI decision-making. That said, many of the same statistical assumptions that are made in trend analyses are also true of level comparisons. For simplicity, refer to page 5 when interpreting the indicators of **baseline instability**, **unanticipated trends**, and **autocorrelation** because that information applies in level comparisons as well.

There are some additional concerns that are worth noting. Unlike trend analyses, level comparisons assume that there are no significant trends in *any* phase. In other words, the effect size estimate is a comparison between two stable conditions. If, for example, there is a significant trend in the B-phase(s), the effect sizes reported in the LevelsResults worksheet may be poor summaries of the intervention's true impact. In the example above, the significant trend from low to high would suggest that the student did not respond immediately to the intervention, so the effect size estimate does not adequately reflect recent outcomes—with time, this estimate is likely to grow.

Also note that although Hedge's *g* and percent change (described below) seem straightforward, both are susceptible to problems. For example, when tracking count data, like the count of main ideas recorded correctly or the number of office disciplinary referrals (ODRs) per week, the distribution can meaningfully deviate from a normal bell curve. Non-normal distributions can also be found when data are percentages or proportions, like the percent of items met on an organization checklist or the proportion of observation intervals when a behavior occurred. Always check the slope, skew, kurtosis, outlier, autocorrelation, and boxplot indicators on the spreadsheet to avoid problems arising from unusual datasets. But even when all indicators appear fine, users must use common sense: If Hedge's *g* and/or percent change do not match the user's visual analysis, consider alternatives.

In the example depicted in Figure 4, the effect size of 1.66 is moderate, so the conclusion would be that the intervention had *some* meaningful impact. In this instance, users might continue the intervention, particularly given the upward trend in the intervention phase (assuming an upward pattern is desired, like in the case of time on-task or proportion of work completion). If the effect size was less than .87, however, the conclusion would be that the intervention was implemented with fidelity. In such instances, the practitioner might drastically revamp the intervention plan or refer for additional services.

It is important to note that effect sizes for single-case design are interpreted differently than between-group designs, with "large effects" requiring estimates greater than 2.67 (see Auerback & Zeitlin, 2014). For easy interpretation, moderate effects are shaded yellow and large effects are shaded green. Confidence intervals (*CI*) for Hedge's *g* are also provided. When this range does not include 0, as is true in the example in Figure 4, the likelihood that the observed change occurred by chance is less than 5%. In other words, it is "statistically significant."

To simplify the results for parents and teachers, a "percent change" estimate is also provided. Users might say, for example, that there was "…a 32% increase in the behavior over previous levels (i.e., baseline)." Percent change is calculated as:

$$100 \times (M_2 - M_1)/M_1$$

WHAT IF THERE ARE PROBLEMS?

The last section of the LevelsResults worksheet provides the Percent of Nonoverlapping Data (PND) and Percentage of Data Exceeding the Median (PEM) methods. Both are "non-overlap," nonparametric effect sizes, which means that fewer assumptions are made about the data (e.g., normal distributions). But the tradeoff is that fewer data are considered. We strongly recommend Hedge's *g* and percent change unless there are clear problems with those estimates (e.g., significant autocorrelation).

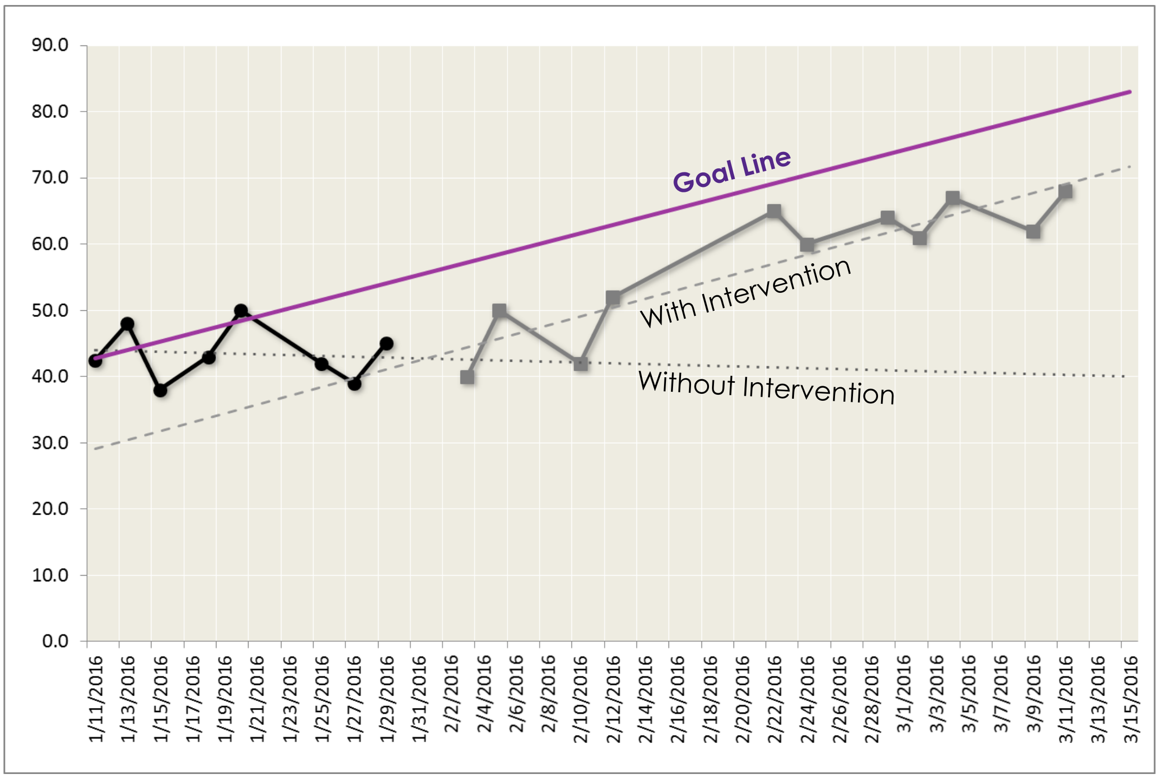
There are additional options when the assumptions of parametric effect sizes have been violated. When in doubt, use a stats package like *R* or online single-case effect size calculator. For up-to-date tools, we recommend the site built by James Pustejovsky and Daniel Swan (2018) available here.

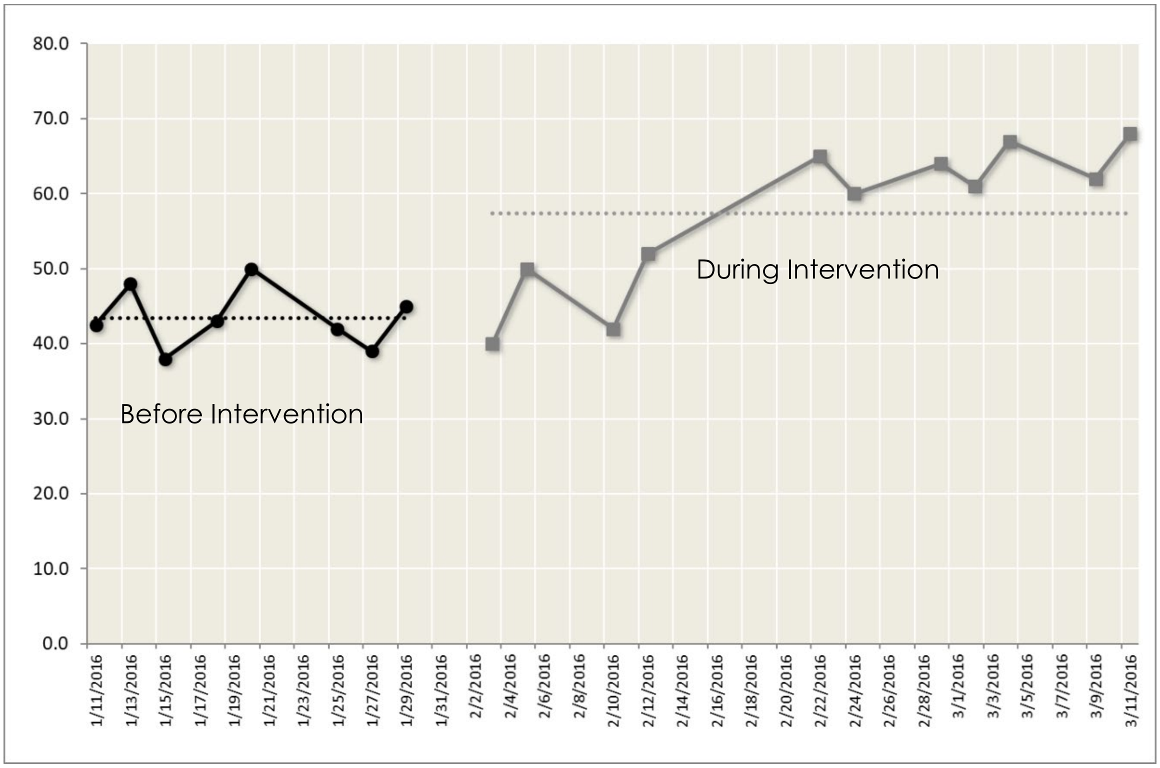
EXPORTING THE GRAPH INTO A REPORT

To copy the graph into a report, click the "Copy Graph" button to the right of either the trend or level graph. Users will see a message box confirming that the image has been saved to the clipboard. Then open the target document and paste **as a picture**. In Word, users might need to select the option of "Paste Special" and then select "Picture" (both PNG and JPEG work well). Then resize as needed.

The advantage of pasting as a picture is that it (a) severs the links between Word and Excel that are sometimes created when cutting-and-pasting between the two programs; and (b) ensures that the axis labels are shown in their entirety (e.g., dates can get truncated when pasting as anything other than a picture). Once the graph is in Word, users can add labels to the graph with text boxes, as depicted in Figure 5. Note that the examples use consumer-responsive labels.

FIGURE 5: GRAPHS EXPORTED AS PICTURE WITH ADDED LABELS




80.0
70.0
60.0
50.0
40.0
30.0
20.0
10.0
0.0
During Intervention
Before Intervention
1/11/2016
1/13/2016
1/15/2016
1/17/2016
1/19/2016
1/21/2016
1/23/2016
1/25/2016
1/27/2016
1/29/2016
1/31/2016
2/2/2016
2/4/2016
2/6/2016
2/8/2016
2/10/2016
2/12/2016
2/14/2016
2/16/2016
2/18/2016
2/20/2016
2/22/2016
2/24/2016
2/26/2016
2/28/2016
3/1/2016
3/3/2016
3/5/2016
3/7/2016
3/9/2016
3/11/2016

References

Auerbach, S., & Zeitlin, W. (2014). *SSD for R: An R Package for Analyzing Single-Subject Data*. New York: Oxford.

Durlak, J.A. (2009). How to select, calculate, and interpret effect sizes. *Journal of Pediatric Psychology*, *34*, 917-928. doi: 10.1093/jpepsy/jsp004

Pustejovsky, J. E. (2018). Using response ratios for meta-analyzing single-case designs with behavioral outcomes. *Journal of School Psychology*, *16*, 99-112. doi: 10.1016/j.jsp.2018.02.003

Pustejovsky, J. E. & Swan, D. M. (2018). Single-case effect size calculator (Version 0.4) Web application. Retrieved from https://jepusto.shinyapps.io/SCD-effect-sizes/

ENDNOTES

[i] Dr. Schultz has created separate "markdown" files for R that walk the user through additional analyses that can be applied to AB designs. The user simply copies the data exported to the .txt file and pastes into the appropriate line of code in the markdown file. Those materials and instructions are available at SchoolPsychologyTech.org