Applying the generalised logistic model in single case designs: modelling treatment-induced shifts

Peter Verboon
Gjalt-Jorn Ygram Peters

Department of Psychology and Educational Sciences,
Open University, The Netherlands

PAPER SUBMITTED FOR PUBLICATION

NOVEMBER 2017

Correspondence concerning this article should be addressed to Peter Verboon, Department of Psychology and Educational Sciences, Open University, P.O. 2960, 6401 DL, Heerlen, the Netherlands, Tel: +31 45 5762786, Email: Peter.Verboon@ou.nl

Applying the generalised logistic model in single case designs: modelling treatment-induced shifts

Single case designs (SCDs) are increasingly recognized as important tools in behaviour modification research and other fields, enabling researchers to model high-resolution changes in psychological or behavioural variables over time. Common approaches to analysis of SCDs are based on comparing means before and after an intervention or modelling the slope of a change using regression-based techniques (Manolov & Moeyaert, 2017a). However, in most situations, assumptions of these approaches are violated because what is modelled is neither a discontinuous shift from one value to another, nor a linear unbounded change over time, but instead a shift from one value to another, where both the initial and the final value are more or less stable over time. Accurate modelling of this shift provides more information about treatment effects. In this paper, we introduce a technique for such modelling. We illustrate this technique at the hand of a data by Singh et al. (2007) and provide a brief tutorial to make these techniques widely accessible.

SCDs are important because they provide a means to determine the effectiveness of interventions at an individual level (Barlow, Nock, & Hersen, 2009). Much methodological research has been devoted to effect size measures in SCD because an accurate effect size supports the development of evidence-based interventions (Parker et al., 2005; Parker, Vannest, & Davis, 2011; Parker & Hagan-Burke, 2007). An effect size can be considered accurate if it provides a reliable indication of for instance the improvement of a patient after or during treatment. The type of effect size is closely related to what type of analysis of SCDs is chosen (Lenz, 2015; Vannest & Ninci, 2015). Two basic classes of analyses can be distinguished: first regression based methods among which also multilevel analysis (Baek et al., 2014), and second non-parametric methods. A recent overview of analysis techniques for SCD is given by Heyvaert and Onghena (2014).

Regression based approaches for analysing SCDs usually consider a linear model. In a pre-post design the piece-wise regression model (Center, Skiba, & Casey, 1985; Huitema & Mckean, 2000) is used to model a linear trend separately for both phases, one before the intervention and one during or after the intervention. This model compares the intercepts and slopes between both phases of the design. Intervention effects are simply derived from the differences in slopes and intercepts.

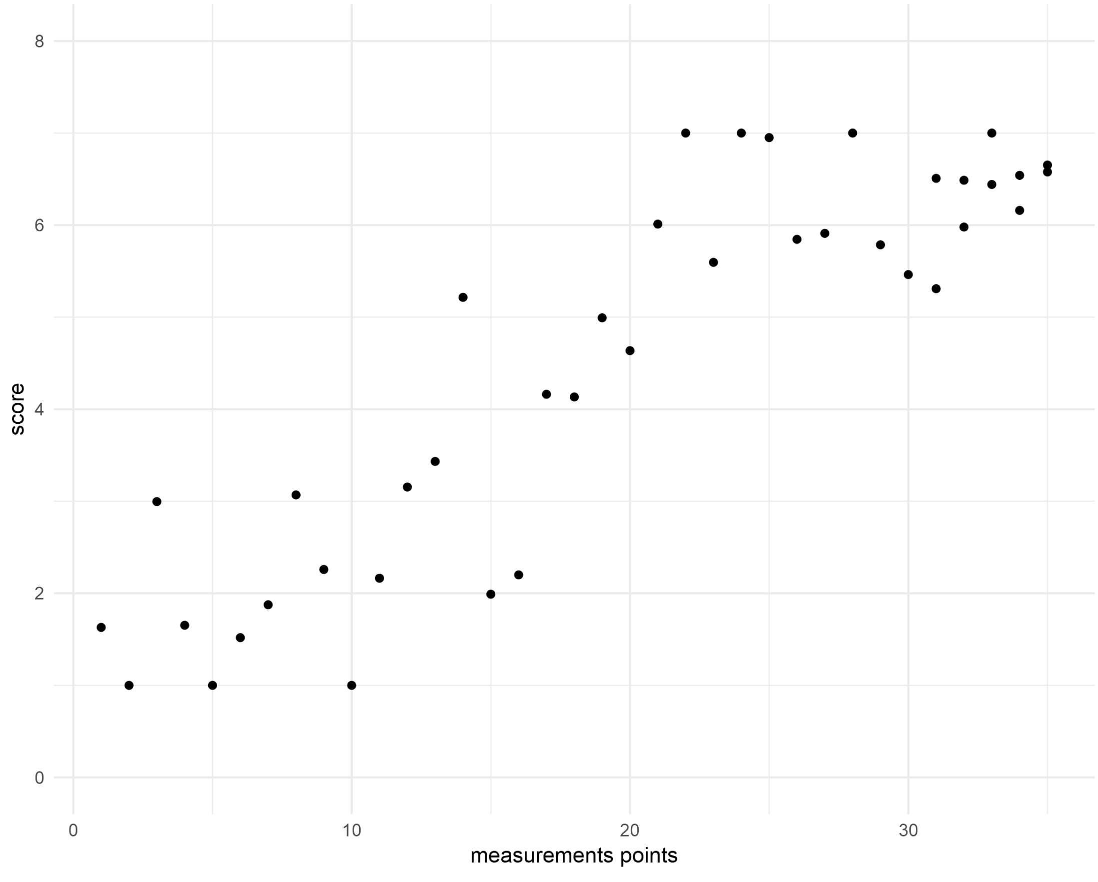
In many situations, for instance when the effects of a therapy are monitored, the criterion is measured by some sort of questionnaire or other instrument that has a limited range. The observed or measured improvement of clients in a therapeutic setting is limited by the scale of the instrument. A seven point Likert type scale is an example of such an instrument. If clients rate how they feel with a maximum score of seven there is no further room for improvement. The score of seven in this example constitutes a ceiling in the therapy effect. Likewise, such an instrument has a floor, which is the minimum value of the scale. A linear model may not be valid when there floor and ceiling effects in the data.

Treatments and interventions in the clinical and health psychology practice are often protocolled. They are designed based on knowledge of the behaviour, cognitions, or affective associations that are targeted. If implemented properly, they will affect the areas of human psychology for which they were designed, thereby improving the target behaviour or condition. However, no psychological theory or combination of theories explains behaviour or psychopathology completely. Therefore, evidence- and theory-based interventions and treatments are necessarily limited in terms of the effect they can have: at most, they can have the maximum achievable effect in all areas they target. This characteristic manifests as a ceiling effect for treatment effectiveness. Such a constraint on effectiveness means that the association between time in treatment and treatment effectiveness cannot be linear.

For data that have a curved character a nonlinear model should provide better estimates of effect sizes than linear regression approaches. In this paper we present a model that enables the estimation of effects in a pre-post SCD design when the data have floor and ceilings. First we present an example of the model and show its mathematical characteristics. Then we show how the model performs compared to other models in a simulation study. Finally we discuss some possibilities for future research.

The problem with ceilings

This distribution shown in Figure 1 is likely to model an ideal intervention process, with the X-axis representing time (e.g. in days) and the Y-axis an outcome (where higher values are more desirable). The first five measurements are before the intervention. The values show measurement error around what is essentially a plateau. Once the intervention commences, however, each session has (on average) some effect to improve the outcome. In this ideal situation, once all targeted areas have been improved, no additional effects can be expected: therefore, after roughly 20 days, the intervention no longer has any effect and another plateau is reached.



[INSERT FIGURE 1 HERE – PROVISIONALLY INCLUDED FOR THE CONVENIENCE OF PEER REVIEW] Figure 1. Example with generated data from generalized logistic model for t=6 to 30 (B = .4, x_0 = 10, v=1). Random data generated for first five and last five points, normally distributed as respectively $N(1.5, 0.5)$ and $N(6.5, 0.2)$.

A simple linear model seems to predict these data rather well, see Figure 2. The deviance (sum of the squared residuals) of the linear model is D_{lm}= 34.3, with R^2 = .80. This is partly due to the fact that the pre-intervention and stability phases are rather short in this example.

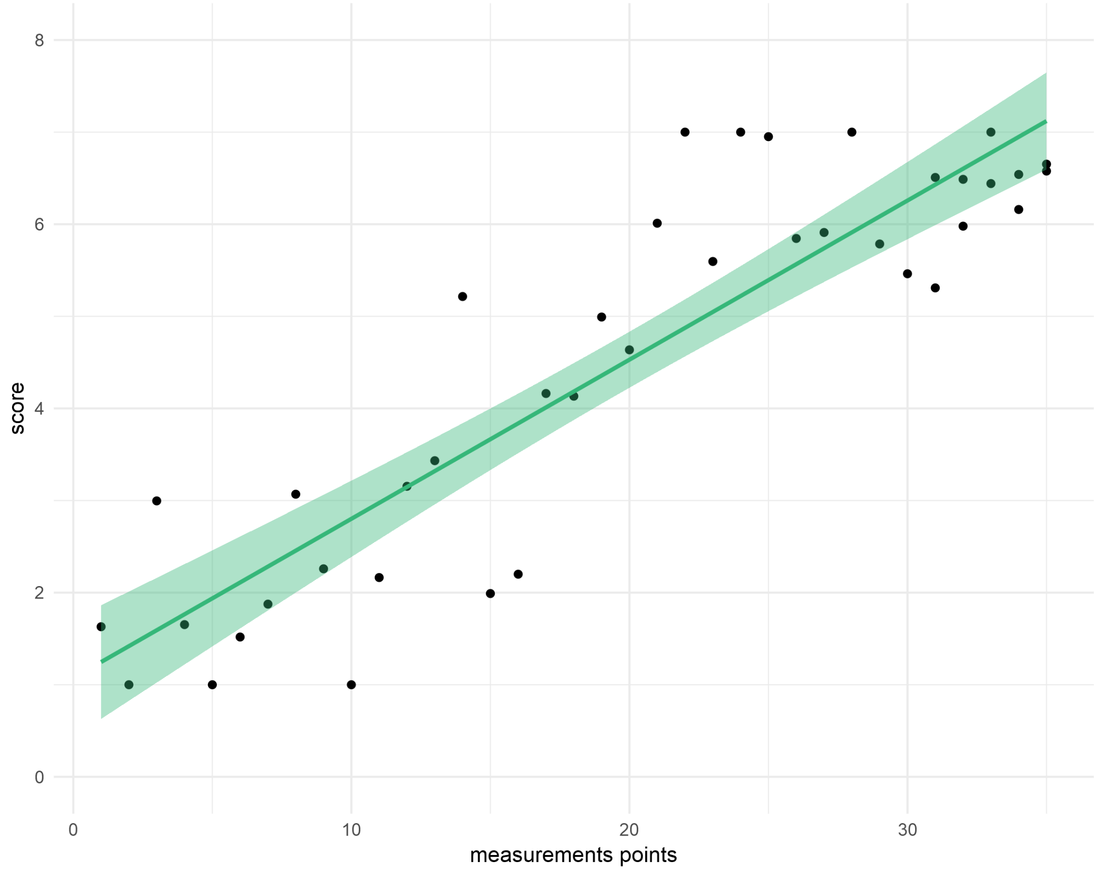


Figure 2. Example with generated data from generalized logistic model with linear fit added (b_0 = 1.07 ; b_1 = .17; R^2 = .80).

However, the residuals from the linear model seem to show a cyclic or auto-correlated pattern, as Figure 3 clearly shows. One of the assumptions for unbiased parameter in linear regression estimates is homogeneity of the residuals and in this example this assumption is violated. This is an indication that the linear model is not the correct model to describe these data.

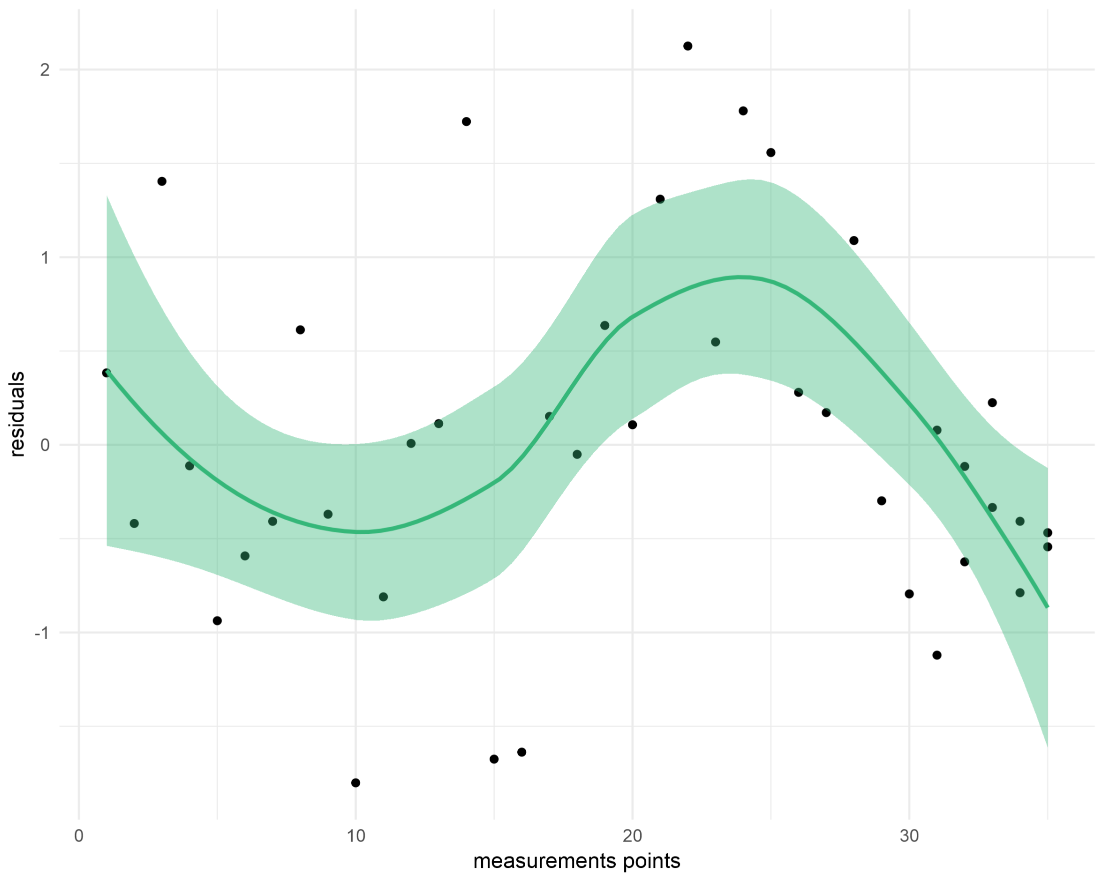


Figure 3. Residuals from linear model of example data.

Despite the high squared correlation the line misses some important information, in particular the strong increase in scores somewhere between the 15^{th} and 20^{th} point.

Consequences for tests of the intervention effect

To test the effect of the intervention, a classic approach is to compare the two means before and after (or during) the intervention. The effect size of the intervention in this approach is Cohen's (Cohen, 1992) *d* or simply the difference between the two means divided by the pooled standard deviation (Rosenthal, 1978). In this example *d* = -1.80, with 95% CI [-2.8, -0.8], a strong effect, which corresponds with the visual inspection of the data.

However, claiming an intervention effect because the means in both phases are different is not correct (Huitema & McKean, 2007). When there appears to be a trend in the data (e.g. scores increase over time, independent of the intervention) simply comparing the means of the outcomes in the two phases may actually lead to wrong conclusions (Center et al., 1985). The trend, instead of an intervention effect, may be responsible for the different means in the two phases. Therefore, it is important to incorporate a trend effect in a research model for SCD data.

To adequately model such trends, a piece-wise regression model (Center et al., 1985; Huitema & Mckean, 2000) can be used. Piece-wise regression models a linear trend separately for both phases. That is, the intercepts and slopes of two regression lines are compared before and after the intervention. See Figure 4 for an illustration. This model is given by:

$$\mathrm{y} = b_0 + b_1\mathrm{D} + b_2t + b_3\mathrm{D}(t - k) + e \tag{1}$$

where y is a vector of length n, n is the total number of measurements, e is a vector with random independent error, D is a dummy which distinguishes the intervention phase (D=1) from the pre-intervention ('control') phase (D=0), and t is the variable representing time. The index t is measured in relevant time-units (e.g. day or week number). In variable k indicates the final measurement in the pre-intervention phase and should be chosen such that the values (t - k) start with 0 in the intervention phase. If t is simply taken as the observation rank number the observations are assumed to be measured at equal time intervals and k represents the number of measurements in phase A (n_{A}), and t runs from 0 to n-1).

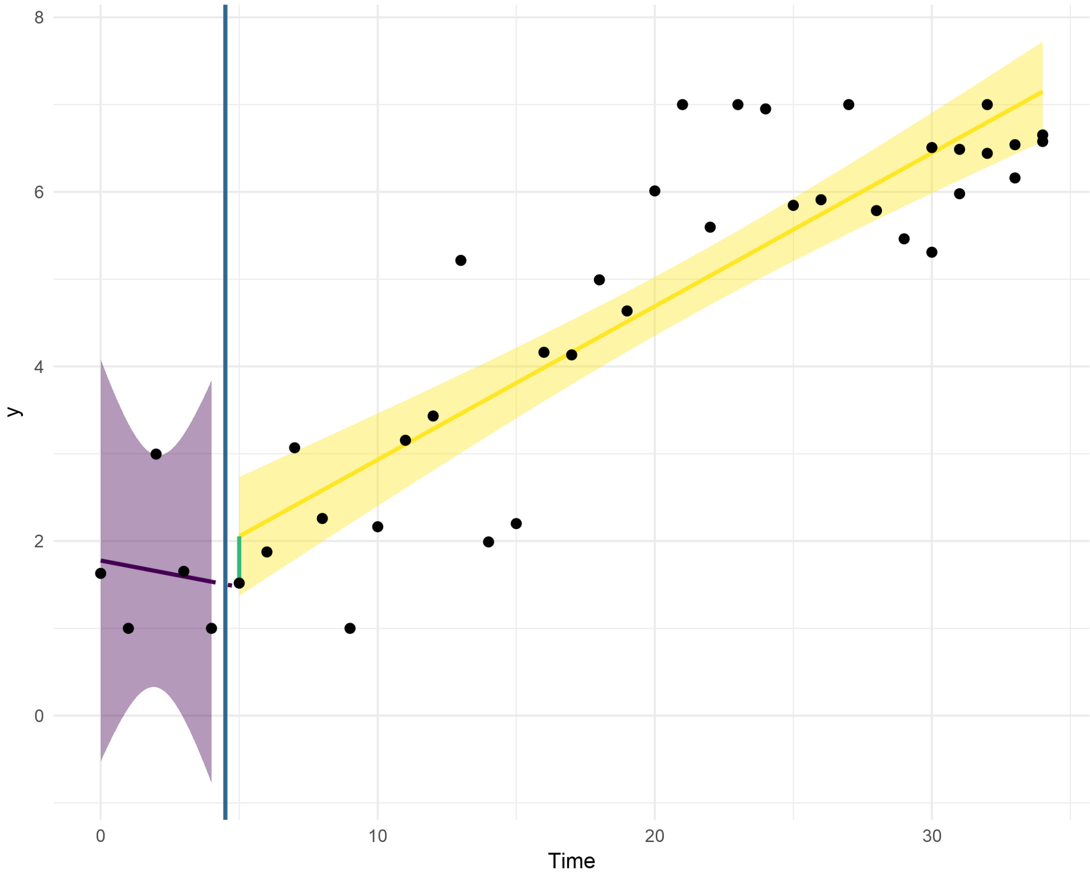


Figure 4. Piecewise regression on example data with trend effect. The blue line indicates (the start of) the intervention. The green line represents the level effect.

In the piece-wise regression (PWR) model b_0 is the score at T=0 (1.8 in this example), b_1 can be interpreted as the change in level between phase A and B, not confounded with possible trend effects. This effect (0.58 with 95% CI [-1.6, 2.7] is represented by the (short) green line; it is the difference between the predicted scores of both regression lines at the first measurement of the second phase. The trend in the baseline phase (A) is captured by b_2 (-0.06 with 95% CI [-0.7, 0.6]) and the *change* in trend from phase A to phase B by b_3, this is 0.24 with 95% CI [-0.4, 0.9]. In this example the interest should be in b_3, the change in slope. The post-intervention line has a slope of about .18. The very large confidence intervals around b_1, b_2 and b_3 are due to very heterogeneous regression intercept and slope in phase A, illustrated by the wide purple zone in Figure 4. Since there are only 5 data points, high levels of heterogeneity can be expected. The deviance of this piecewise model is D_{pw} = 33.7, which is slightly better than the linear model.

When only a level effect is present in the data, such as shown in another example in Figure 5, the b_1 (4.18) parameter would be of primary interest. The change from the slope in the pre-intervention phase (b_2 = -0.04) to the flat line in the post-intervention phase is as expected 0.05 (b_3). Cohen's *d* is -7.9 with 95% CI [-10.5, -5.2] in this example.

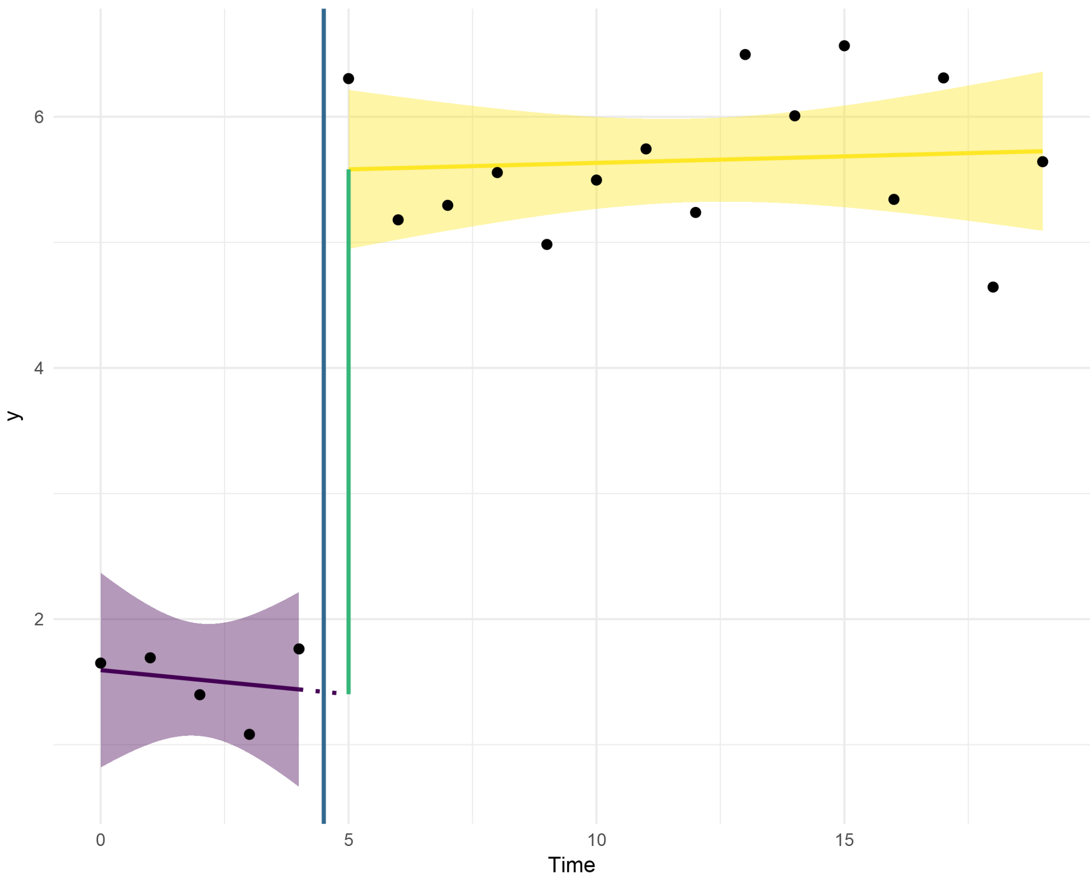


Figure 5. Piecewise regression on example data with phase effect. The blue line indicates (the start of) the intervention. The green line represents the level effect.

For this piecewise regression method the following effect size is defined (Parker & Brossart, 2003):

$$ES_{pwr} = (R_m^2 - R_0^2)/(1 - R_0^2), \tag{2}$$

where R_m^2 represents the squared multiple correlation coefficient of the piecewise regression model and R_0^2 the squared multiple correlation coefficient of a model with intercept and trend parameters only. This latter "null" model ignores the phase differences, so the resulting effect size can be viewed as the explained variance in the dependent variable unaccounted for by the null model. Parker and Brossart (Parker & Brossart, 2003) warn that classical guidelines concerning the strength of effect sizes are not valid for new analytic techniques like the ones that are suggested for SCD. This means that effect sizes are primarily useful for comparison between studies with similar designs and for accumulation.

In many situations, it is not only important to know that there exist an effect and how strong it is, but also at what point in time the improvement due to the intervention started and when the improvement stabilized. For this type of questions it is better to fit a curve to the data that has the form of a sigmoid function, because it is more flexible and seems to be better suited to model the empirical process.

The generalized logistic model

A sigmoid function can be defined in many ways. Here we choose the generalised logistic function (GL), which is defined as follows:

$$y(t) = A_B + \frac{(A_T - A_B)}{(1 + e^{-B(t-t_0)})^{\frac{1}{v}}} \tag{3}$$

This model has the advantage that it is parametrized relatively straightforwardly: the analysis estimates the initial plateau and the post-intervention plateau as well as when the change starts and stops. Specifically, the variable y(t) is the outcome at moment t (t = 1, … , n). The parameters A_B and A_T are the asymptotes that indicate, respectively, the minimum (floor) and maximum (ceiling) of the curve. The parameter B is the growth rate, indicating how steep the curve is. The parameter v indicates near which asymptote the maximum growth occurs and t_0 corresponds to the time point at which the curve is at its midpoint (when $v = 1$). For the parameter values: $A_B = 0$, $A_T = 1$, $B = 1$, $v = 1$, and $t_0 = 0$, this function simplifies to the well-known logistic function.

The generalized logistic function was fitted on the example data (see Figure 6) with the *R* function *nlsLM()* from the package *minpack.lm* (Elzho, Mullen, Spiess, & Bolker, 2016). The resulting curve fitted the data well: R^2 = .85 and the deviance D_{gl} = 26.3, which indicates as expected a better fit to the data. The parameters obtained from this analysis were t_0 = 17.0, B = 0.20, A_B = 1.2, A_T = 7.0 and v = 1 (fixed). From this analysis we learn that the process starts at 1.2 and ends at 7.0.

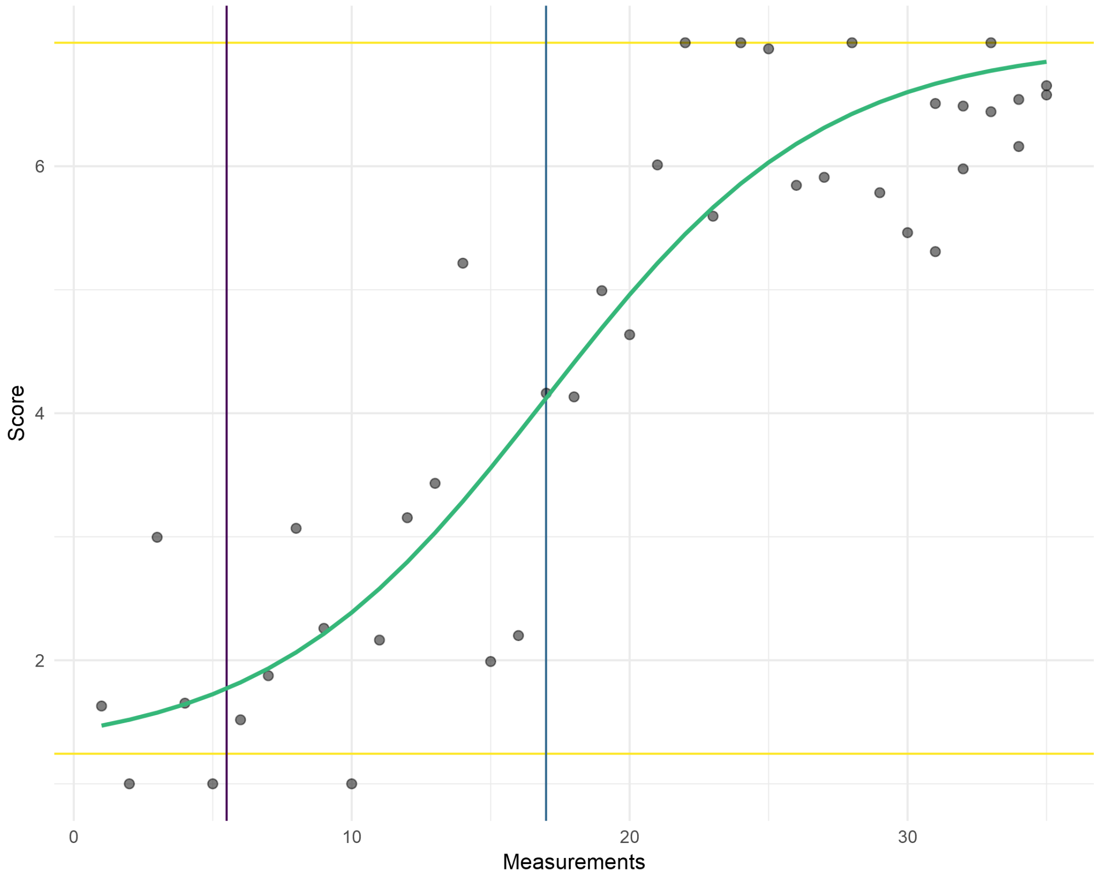


Figure 6. Generalized logistic function fitted to the example data.

At about measurement 17 (12 measurements after the intervention started) the rate of increase in scores is largest. The growth rate is 0.2.

A general effect size could be defined in line with Cohen's *d* as:

$$ES1_{gl} = (A_T - A_B) / SD(y). \tag{4}$$

where SD(y) is the SD of y from a particular subject. Instead of the means in both phases the estimated floor and ceiling are taken in this formula. For this example the $ES1_{gl}$ = 2.75. Alternatively the theoretical or empirical range of the scale of the measurement instrument could be used in the denominator, as:

$$ES2_{gl} = (A_T - A_B) / \text{Range}(y). \qquad (5)$$

This ES indicates the proportion of the scale that is improved according to the floor and ceiling of the fit function, here $ES1_{gl}$ = .96. The growth rate parameter can also be viewed as measure of effect size. An example of 6 different growth rates is shown in figure 7. It does not indicate how large an effect is, but how fast the effect is reached. From a practical perspective it is conceivable that a smaller effect (as measured by ES_{gl}) that is reached relatively quickly is preferable over a larger effect that takes a long time to be reached.

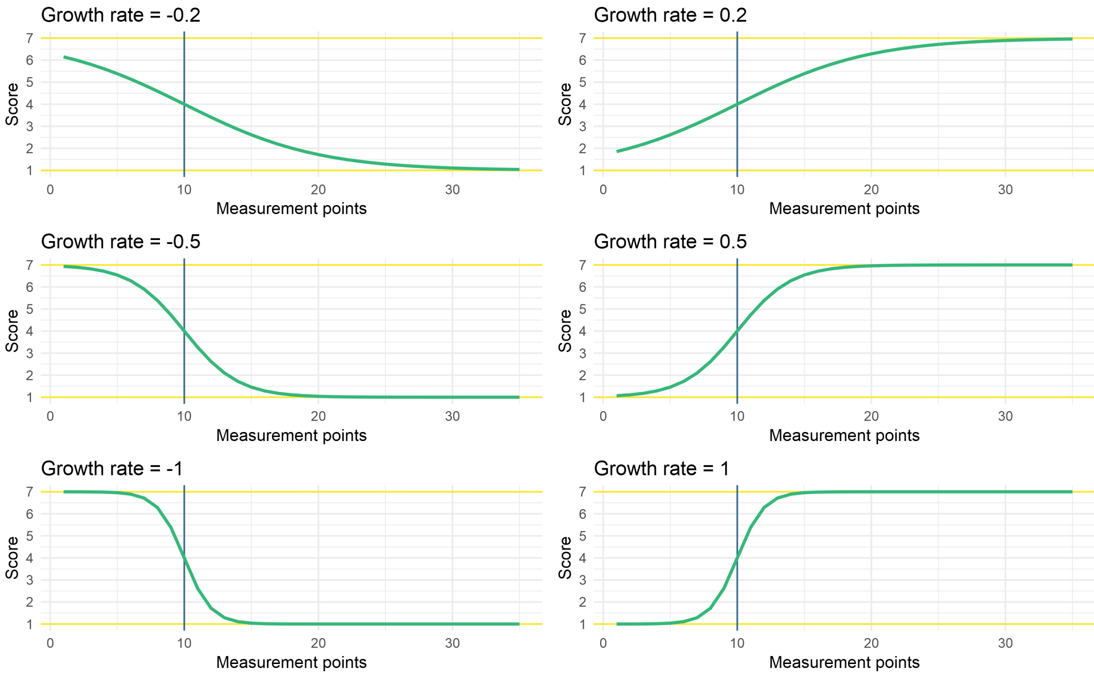


Figure 7. Examples of growth rates for fixed v=1, x_0=10, bottom=1 and ceiling=7.

The function *genlog*() has been build around the optimizing function *nlsLM()* to run the GL model with sensible starting values and minimum and maximum constraints for the parameters (see Appendix for a small tutorial, where function *piecewiseRegr()* is also explained). Sensible starting values and constraints are necessary to avoid convergence problems of the algorithm. The *genlog*() function also contains the option to plot the result (e.g. Figure 6) using the *ggplot2* package (Wickham, 2009).

The A_T is constrained around the maximum value of the scores of the dependent variable: [max(y) - 2, max(y) + 1], A_B is constrained around the minimum value of the scores: [min(y) - 1, min(y) + 2]. The growth parameter is constrained between -2 and +2. Finally, the inflection point (t_0) is constrained between the 2 units after the first measurement and ten units before the last measurement.

Default starting values for the parameters are: for $t_0 = n_A+4$, for A_B=min(y) , for A_T = max(y), and for B = 0. All of the constraints and starting values can easily be changed if the data require other values.

Empirical example

In their extensive review paper about SCD and methodologies to analyse them Manolov and Moeyaert (Manolov & Moeyaert, 2017a) analysed a dataset from Singh et al. (2007). In the present paper we will also use these data to illustrate the GL model and compare the results with those presented in the Manolov and Moeyaert paper. The data were obtained from three individuals measuring their verbal and physical aggression before and after an intervention, which consisted of mindfulness training for controlling aggressive behaviour. The individuals were diagnosed with several mental disorders such as depression, schizoaffective disorder, borderline personality, and antisocial personality. These data are considered representative for single case data in the literature (Shadish & Sullivan, 2011).

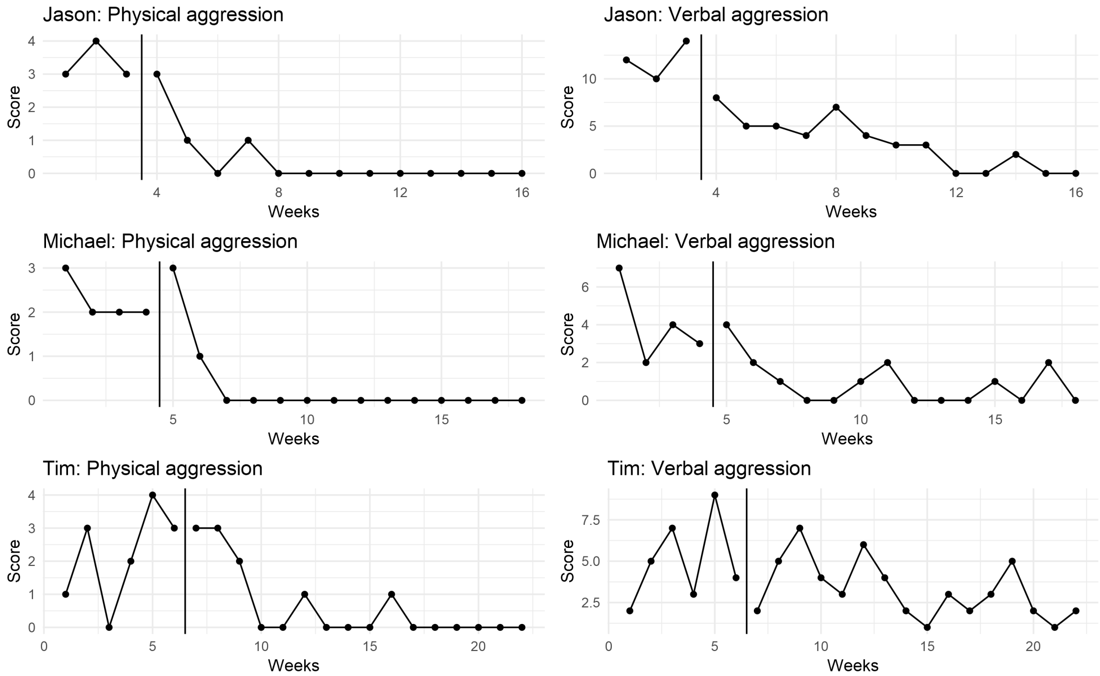


Figure 8. Representation of the 6 datasets obtained from Singh et al. (2007). The vertical grey line distinguishes the pre- and post-intervention phase.

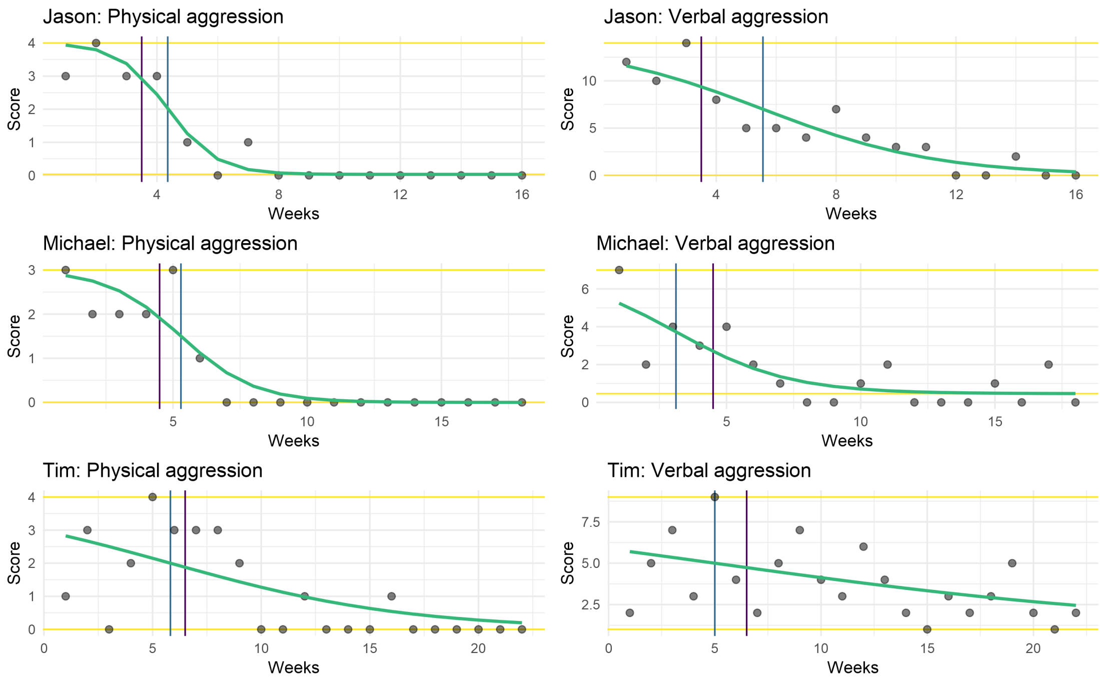


Figure 9. Data from Singh et al. (2007) analysed with the GL model. The vertical blue line distinguishes the pre- and post-intervention phase. The vertical purple line indicated the inflection point (t_0) and the yellow horizontal lines indicate the floor and ceiling values.

In table 1 the results are presented of the three effect size statistics obtained in the GL analyses and compared with the piecewise regression analysis (PWR) and Cohen's *d* (see also Manolov and Moeyaert (2017). Cohen's *d* only compares the level effect between the two phases, PWR compares both level and linear trend, and GL fits a curved effect.

Table 1. Comparison of results GL with PWR and Cohen's d.

	Level	GL					PWR	
	Cohen's *d*	R^2	Growth rate	ES1_gl	ES2_gl	ES_pwr	Level change	Trend change
Physical aggression								
Jason	0.87	.92	-1.24	2.76	0.99	.42	-2.10	-0.14
Michael	3.53	.85	-0.74	2.66	0.75	.19	-0.50	+0.19
Tim	2.53	.39	-0.16	2.94	1.00	.26	-1.85	-0.59
Verbal aggression								
Jason	3.38	.76	-0.67	2.51	0.79	.52	-7.15	-1.62
Michael	2.14	.67	-0.47	3.42	0.47	.25	+0.27	+0.87
Tim	0.87	.19	-0.09	3.77	0.57	.11	-2.20	-0.69

GL: Generalized logistic regression model
PWR: Piecewise regression model

From visual inspection we learn that Jason has made the biggest improvement, both with respect to verbal and physical aggression. However, this effect is based on only three measurements in the first phase. Despite the low number of data points in the first phase the effect in Jason's data is well captured by B (growth rate), $ES2_{gl}$ and ES_{pwr}. Furthermore, the R^2 indicates that the GL model can very well summarize Jason's data. $ES1_{gl}$ is less convincing: although its values are larger than for ES2, they are relatively small compared to the ES2 of the two other persons. In fact, for verbal aggression, it has a lower value for Jason than for the other two, which is caused by the larger variance of Jason's data (starting with values around 12, which is much larger than for the other two). Cohen's *d* obviously suffers from the same problem.

Results from GL of Tim's physical aggression data also show large ES. However, the R^2 is low and also the growth rate is much less than that of the other persons. In all cases the floor and ceiling effect are in line with what should be expected when we visually inspect the data.

For valid interpretation of the GL results we recommend to first inspect the R^2. If the R^2 is low, the curve cannot fit the data well and all ES values are rather meaningless. Keep in mind that in SCD, the R^2 values are usually larger than in “classical” regression situations with large N, since there are a limited number of data points in SCD.

Discussion

The present paper discusses a new method to analyse experimental single case data based on a generalized logistic model. The underlying assumption of this method is that intervention effects represent the shift of an individual’s scores from one plateau to another, and that the individual’s scores are limited by floor and ceiling effects, which are caused by the measurement instrument and by natural limits of the process under study. This implies that the linear models to estimate the intervention effect are at best suboptimal because their assumptions are violated, and, relatedly, they fit the data poorly. The generalized logistic model seems better equipped to deal with these floor and ceiling effects of the measurement instruments.

To test the proposed method we built the R function *genlog* around a general existing optimizing function, with this new function providing sensible default starting values and constraints. Running the *genlog* function yields parameter estimates and also provides visualisation of the data and the fitted function. Together with the function we proposed two simple effect size measures derived from Cohen’s D. In addition, we argued that the growth parameter of the function could serve as an additional effect size measure, indicating the speed of the intervening process. How to qualify the effect sized we proposed as large or small is a question that remains to be addressed (see also Manolov, Gast, Perdices, & Evans, 2014). Visual inspection of the data was used here to gauge the plausibility of our effect sizes. More studies are necessary to obtain a better understanding of these effect sizes.

Based on a well-known single case data (Singh et al., 2007) we illustrated the generalized logistic model. The Singh data are also discussed in Manolov and Moeyaert (2017a) and used to compare a wide variety of single case methods. The model was applied to these data and compared with the piece-wise regression model. The generalized logistic model provided sensible outcomes that seem to add to the understanding of the intervention process. Based on these analyses, we recommend that one should combine the result of the model fit with that of the estimated growth parameter and the second effect size, which is based on the range of the data, to obtain informative outcomes.

Due to the character of single case data, the parameter estimates of the generalized logistic model are not robust in the sense that they depend on parameter constraints and starting values. With relatively few data points and four parameters to estimate this is logical consequence. Fixing the top and ceiling values after visual inspection can easily improve the robustness of the remaining parameters.

With multiple single case data (i.e. replicated n-of-1 designs), future research should focus on whether this model can be incorporated in a multilevel context. In Baek et al. (2014) the integration of single case results by multilevel analyses is discussed. It is shown by these authors how the piecewise regression model can be incorporated in a multilevel framework. Moeyaert et al. (2014) found empirical evidence that the fixed effects in three level analyses of single case studies are unbiased, a result that was found earlier in two level analysis (Ferron, Bell, Hess, Rendina-Gobioff, & Hibbard, 2009). It was also found by combining more than 30 studies that the mean squared error was hardly influenced by the small single case designs. It is expected that this finding generalizes to the model we have proposed in the present study. Combining many studies has the additional advantage that the estimated model parameters will show more robustness (i.e. be less dependent on the starting values).

In this paper, we have presented another tool to add to the already wide collection of SCD approaches (Manolov & Moeyaert, 2017b). It is based on the idea that most effects of interventions have a natural limit. Based on this simple premise we have proposed a model that would represent this idea. The software we have presented is Free and Open Source Software, implemented in the popular statistical environment R, and easy to apply, with some additional support in a short tutorial.

References

Baek, E. K., Moeyaert, M., Petit-Bois, M., Beretvas, S. N., Van den Noortgate, W., & Ferron, J. M. (2014). The use of multilevel analysis for integrating single-case experimental design results within a study and across studies. *Neuropsychological Rehabilitation*, *24*(3–4), 590–606. https://doi.org/10.1080/09602011.2013.835740

Barlow, D. H., Nock, M. K., & Hersen, M. (2009). *Single case experimental designs. Strategies for studying behaviour change* (3rd ed.). Boston, MA: Allyn and Bacon.

Center, B. A., Skiba, R. J., & Casey, A. (1985). A methodology for the quantitative synthesis of intra-subject design research. *The Journal of Special Education*, *19*(4), 387–400.

Cohen, J. (1992). A power primer. *Psychological Bulletin*, *112*(1), 155–159.

Elzho, T. V., Mullen, K. M., Spiess, A.-N., & Bolker, B. (2016). minpack.lm: R Interface to the Levenberg-Marquardt Nonlinear Least-Squares Algorithm Found in MINPACK, Plus Support for Bounds.

Ferron, J. M., Bell, B. A., Hess, M. F., Rendina-Gobioff, G., & Hibbard, S. T. (2009). Making treatment effect inferences frommultiple-baseline data: The utility of multilevelmodeling approaches. *Behavior Research Methods*, *41*, 372–384. https://doi.org/https://doi.org/10.3758/BRM.41.2.372

Heyvaert, M., & Onghena, P. (2014). Randomization tests for single-case experiments: State of the art, state of the science, and state of the application. *Journal of Contextual Behavioral Science*, *3*(1). https://doi.org/10.1016/j.jcbs.2013.10.002

Huitema, B. E., & Mckean, J. W. (2000). Design Specification Issues in Time-Series Intervention Models. *Educational and Psychological Measurement*, *60*(1), 38–58. https://doi.org/10.1177/00131640021970358

Huitema, B. E., & McKean, J. W. (2007). Identifying Autocorrelation Generated by Various Error Time-Series Regression Designs. *Educational and Psychological Measurement*, *67*, 447–459.

Lenz, a. S. (2015). Using Single-Case Research Designs to Demonstrate Evidence for Counseling Practices. *Journal of Counseling & Development*, *93*(4), 387–393. https://doi.org/10.1002/jcad.12036

Manolov, R., Gast, D. L., Perdices, M., & Evans, J. J. (2014). Single-case experimental designs: reflections on conduct and analysis. *Neuropsychological Rehabilitation*, *24*(3–4), 634–60. https://doi.org/10.1080/09602011.2014.903199

Manolov, R., & Moeyaert, M. (2017a). How Can Single-Case Data Be Analyzed? Software Resources, Tutorial, and Reflections on Analysis. *Behavior Modification*, *41*(2), 179–228. https://doi.org/10.1177/0145445516664307

Manolov, R., & Moeyaert, M. (2017b). Recommendations for choosing single-case data analytical techniques. *Behavior Therapy*, *48*, 97–114. https://doi.org/10.1016/j.beth.2016.04.008

Moeyaert, M., Ugille, M., Ferron, J. M., Beretvas, S. N., & Van den Noortgate, W. (2014). Three-Level Analysis of Single-Case Experimental Data: Empirical Validation. *The Journal of Experimental Education*, *82*(1), 1–21. https://doi.org/10.1080/00220973.2012.745470

Parker, R. I., & Brossart, D. F. (2003). Evaluating Single-Case Research Data: A Comparison of Seven Statistical Methods. *Behavior Therapy*, *34*(2), 189–211.

https://doi.org/10.1016/S0005-7894(03)80013-8
Parker, R. I., Brossart, D. R., Vannest, K. J., Long, J. R., De-alba, R. G., Baugh, F. G., & Sullivan, J. R. (2005). Effect Sizes in Single Case Research : How Large is Large ? *School Psychology Review*, *34*(l), 116–132.
Parker, R. I., & Hagan-Burke, S. (2007). Useful effect size interpretations for single case research. *Behavior Therapy*, *38*(1), 95–105. https://doi.org/10.1016/j.beth.2006.05.002
Parker, R. I., Vannest, K. J., & Davis, J. L. (2011). Effect Size in Single-Case Research: A Review of Nine Nonoverlap Techniques. *Behavior Modification*, *35*(4), 303–322. https://doi.org/10.1177/0145445511399147
Rosenthal, R. (1978). Combining results of independent studies. *Psychological Bulletin*, *85*, 185–193.
Shadish, W. R., & Sullivan, K. J. (2011). Characteristics of single-case designs used to assess intervention effects in 2008. *Behavior Research Methods, 43*, 971–980.
Singh, N. N., Lancioni, G. E., Winton, A. S. W., Adkins, A. D., Wahler, R. G., Sabaawi, M., & Singh, J. (2007). Individuals with Mental Illness Can Control their Aggressive Behavior Through Mindfulness Training. *Behavior Modification*, *31*(3), 313–328. https://doi.org/10.1177/0145445506293585
Vannest, K. J., & Ninci, J. (2015). Evaluating Intervention Effects in Single-Case Research Designs. *Journal of Counseling & Development*, *93*(4), 403–411. https://doi.org/10.1002/jcad.12038
Wickham, H. (2009). *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York. Retrieved from http://ggplot2.org