Percentage of nonoverlapping corrected data

Rumen Manolov and Antonio Solanas

Department of Behavioral Sciences Methods,

Faculty of Psychology,

University of Barcelona

MAILING ADDRESS

Correspondence concerning this article should be addressed to Rumen Manolov, Departament de Metodologia de les Ciències del Comportament, Facultat de Psicologia, Universitat de Barcelona, Passeig de la Vall d'Hebron, 171, 08035-Barcelona, Spain. Phone number: +34933125844. Fax: +34934021359. Electronic mail may be sent to Rumen Manolov at rrumenov13@ub.edu.

AUTHORS' NOTE

This research was supported by the *Comissionat per a Universitats i Recerca del Departament d'Innovació, Universitats i Empresa* of the *Generalitat de Catalunya* and the European Social Fund.

ACKNOWLEDGEMENTS

The authors would like to thank the anonymous reviewers for their useful comments and suggestions, which contributed to improving the manuscript.

RUNNING HEAD

Percentage of nonoverlapping corrected data

ABSTRACT

The present study proposes a modification in one of the most frequently applied effect size procedures in single-case data analysis – the percent of nonoverlapping data. In contrast to other techniques, the calculus and interpretation of this procedure is straightforward and it can be easily complemented by visual inspection of the graphed data. Although the percent of nonoverlapping data has been found to perform reasonably well in N = 1 data, the magnitude of effect estimates it yields can be distorted by trend and autocorrelation. Therefore, the data correction procedure focuses on removing the baseline trend from data prior to estimating the change produced in the behavior due to intervention. A simulation study is carried out in order to compare the original and the modified procedures in several experimental conditions. The results suggest that the new proposal is unaffected by trend and autocorrelation and can be used in case of unstable baselines and sequentially related measurements.

Key words: percent of nonoverlapping data, single-case designs, effect size, autocorrelation, trend

Single-case designs are useful for obtaining scientific evidence about intervention effectiveness in different behavioral fields of knowledge (Crane, 1985; Gedo, 2000; Tervo, Estrem, Bryson-Brockman, & Symons, 2003). Recent methodological research on single-case data analysis has centered on effect size measures instead of on statistical techniques yielding exclusively *p*-values. This might be due to the recommendations for reporting studies' results (Wilkinson & The Task Force on Statistical Inference, 1999) based on the advantages of effect sizes over statistical significance, such as the focus on the strength of relationship between the intervention and behavior of interest, the possibility to establish different degrees of treatment effectiveness and the avoidance of the sample size dependence (Cohen, 1990; 1994; Kirk, 1996; Kromrey & Foster-Johnson, 1996; Rosnow & Rosenthal, 1989). The importance of effect size measurements in single-case designs has been reflected in the increased amount of recent publications answering the need of evidence-based interventions in the behavioral sciences (Jenson, Clark, Kircher, & Kristjansson, 2007; Schlosser & Sigafoos, 2008; Shadish, Rindskopf, & Hedges, 2008).

From the perspective of an applied researcher in clinical, educational or social settings, a potentially useful effect size index needs to meet several criteria: 1) to perform well in short data series, producing low estimates in absence of treatment effect and higher ones in its presence; 2) to be easy to interpret in applied rather than in statistical terms; 3) related to the previous, it is desirable that the procedure is designed specifically for N = 1 data in order to avoid interpretations based on group designs terminology; 4) to be simple to compute, not requiring expertise, commercial statistical software packages or excessive amount of time; 5) to be easily complemented by visual inspection considering its utility (Parker, Cryer, & Byrns, 2006) and its frequent application (Kratochwill & Brody, 1978; Parker & Brossart, 2003).

As regards the first criterion mentioned, several regression-based techniques have been found to have unacceptable statistical properties (Beretvas & Chung, 2008; Manolov & Solanas, 2008; Parker & Brossart, 2003). These prcedures also require a greater amount of knowledge and of calculus in comparison to the indices related to visual analysis proposed. Considering these latter procedures, Ma's (2006) percentage of data points exceeding the median and Parker, Hagan-Burke, and Vannest's (2007) percentage of all non-overlapping data (PAND) were designed to improve the performance of the percent of nonoverlapping data (PND; Scruggs, Mastropieri, & Casto, 1987), but it has been shown that this is not always the case (Manolov, Solanas, & Leiva, in press). Additionally, the magnitude of effect estimate produced by PAND has a less straightforward interpretation, whereas the Pearson's Phi^2 which can be obtained out of it requires several steps in different software (Schneider, Godlstein, & Parker, 2008).

Taking into account these considerations, the percent of nonoverlapping data (PND) which was designed for single-case data can be regarded as a procedure performing well (i.e., better than its most similar alternatives, although not optimally), being simple to interpret and to compute and closely related to visual inspection. In recent studies, PND has been the most frequently applied procedure to quantify treatment effectiveness in single-case studies and also in meta-analyses (Schlosser, Lee, & Wendt, 2008). Nevertheless, despite its attractiveness to psychologists, PND is not a trouble-free procedure (Allison & Gorman, 1994; Manolov & Solanas, 2008). Therefore, the main objective of the present investigation is to propose a modification of the PND procedure intended to overcome some of its limitations. The performance of the modified index is tested in the context of data sets with different characteristics such as presence or absence of confounding variables (i.e., trend, serial dependence) and of intervention effects. In order to contrast the percentages obtained against known data attributes, Monte Carlo methods were used to construct the data series.

Overcoming the drawbacks of PND

The present study proposes a data correction procedure to be implemented prior to applying the PND. The main aim of the procedure is to eliminate from data a possible preexisting trend not related to the introduction of the intervention. Since the proposal is basically a modification of PND adding an initial data correction step, we refer to the procedure as the Percentage of nonoverlapping corrected data (PNCD). Before a treatment is introduced (i.e., in an AB design's initial phase) it can be reasonably assumed that the behavior of the individual (*y*) or group studied is randomly fluctuating around a certain value, that is, $y_t = \varepsilon_t$. If there is a trend in the behavior, then $y_t = \beta \cdot t + \varepsilon_t$, where β is the trend coefficient (equal to zero in absence of trend) and t is the value of the time variable. The original phase A consists of n_A data points, which when differenced, lead to a new series of $n_A - 1$ values: $\Delta y_{t+1} = y_{t+1} - y_t$. In case there is trend in data $\Delta y_{t+1} = [\beta \cdot (t+1) + \varepsilon_{t+1}] - [\beta \cdot t + \varepsilon_t] = \beta \cdot t + \beta + \varepsilon_{t+1} - \beta \cdot t - \varepsilon_i = \beta + \varepsilon_{t+1} - \varepsilon_t$. ε_{t+1} and ε_t are supposed to be independent and randomly and identically distributed, their mathematical expectancy is assumed to be zero. Given that $E\left[\beta + \varepsilon_{t+1} - \varepsilon_t\right] = \beta$, an estimate of β can be obtained averaging the differenced data series, that is $\overline{\Delta y}$ is used as β. After the trend in the baseline phase is estimated, the whole series (both phase A and B) can be corrected subtracting $\beta \bullet t$ (the trend estimate multiplied by the measurement time) from the original data points. This operation is expected to remove trend from data and, thus, avoid inflation in the percentages obtained by means of PND. Trend is not estimated from the whole data series, since a change in level between the phases may be confounded for trend and such a correction may remove intervention effect. The steps necessary for computing both PND and PNCD are illustrated in a following section. Additionally, R codes were developed for computing both indices and are presented in the Appendices I and II for interventions aiming to increase and decrease the response rate, respectively.

As regards autocorrelation, a difference needs to be established between positive serial dependence and negative one. Higher degrees of positive autocorrelation can be represented by upward or downward trends and, therefore, it can be conjectured that a correction focusing on trend may also have influence on it and attenuate its impact on the effect size index. Negative autocorrelation, however, is related to alternations of dissimilar measurements. It this case the effect of the correction procedure proposed cannot be foreseen and needs to be explored.

Outliers represent another data feature that can distort the magnitude of effect estimates provided by PND. For instance, a single extremely high value in phase A can mask a behavioral change taking place after the treatment is introduced. Outliers can be detected using statistical calculi and can be controlled by means of elimination, winsorization, etc. However, it has to be taken into account that in a single-case study the applied researcher possesses a thorough knowledge of the client and is able to identify which measurement is an extreme and potentially anomalous one and interpret it (e.g., seek for its reason) from a clinical, educational, social, etc. point of view. Such a theoretical interpretation may be more meaningful than an arbitrary statistical treatment of the unexpected datum.

Method

AB series' lengths

Short data series ($N = n_A + n_B$) were included in the present study, since those are more feasible in applied settings: a) $N = 10$ with $n_A = n_B = 5$; b) $N = 15$

with $n_A = 5$; $n_B = 10$; c) N = 15 with $n_A = 7$; $n_B = 8$; d) N = 20 with $n_A = n_B = 10$; e) N = 30 with $n_A = n_B = 15$; and f) N = 40 with $n_A = n_B = 20$.

Data generation

For each combination of n_A and n_B data were according to the model proposed by Huitema and McKean (2000; 2007a):

$$y_t = \beta_0 + \beta_1 \cdot T_t + \beta_2 \cdot D_t + \beta_3 \cdot SC_t + \varepsilon_t,$$

where y_t is the value of the dependent variable at moment t, β_0 is intercept set to zero, β_1, β_2, and β_3 are the coefficients associated with trend, level change, and slope change, respectively, T_t is the value of the time variable at moment t (taking values from 1 to N), D_t is a dummy variable for level change (equal to 0 for phase A and to 1 for phase B), SC_t is the value of the slope change variable being equal to 0 for phase A, and taking values from 0 to (n_B − 1) for phase B, and ε_t is the error term.

The error term (ε_t) was generated following two different models. The commonly used first-order autoregressive model $\varepsilon_t = \varphi_1 \cdot \varepsilon_{t-1} + u_t$, with φ_1 ranging from –.9 to .9 in steps of .1. Since there is evidence that other models, especially a first-order moving average, can be used to represent behavioral data (Harrop & Velicer, 1985), the MA(1) model $\varepsilon_t = u_t - \theta_1 \cdot u_{t-1}$ presented in McCleary and Hay (1980) was studied using 19 values of θ_1: −.9(.1).9. According to the formula $\varphi_1 = -\theta_1/(1 + \theta_1^2)$, this meant that the degrees of autocorrelation ranged from −.4972 to .4972.

For both models the random variable u_t was generated following N(0,1) and, additionally, an exponential and a uniform distribution with the same mean and standard deviation, since normal distribution are not always appropriate models for behavioral measurements (Bradley, 1977; Micceri, 1989). The abovementioned distributions are relevant since they differ in terms of skewness and kurtosis from the Gaussian distribution.

The values of β_1, β_2, and β_3 (.06, .3, and .15, respectively) were chosen by trial and error, a procedure also followed by Parker and Brossart (2003) and Brossart, Parker, Olson, and Mahadevan (2006), aiming to avoid floor and ceiling effects in the percentages obtained (Manolov & Solanas, 2008). In addition, the values of those coefficients were determined in a way to produce equivalent mean shifts in the case of trend, change in slope, and change in level for $n_A = n_B = 5$ data series. In any case, the specific beta-values are not essential, since they only serve to construct data series with and without trend or intervention effect and, thus, create a common background for comparing PND and PNCD.

Analysis

Prior to presenting in detail the steps needed to carry out the two effect size procedures included in the present study, an example of a fictitious data set is presented. Consider a psychological single-case study educating parent to interact with children diagnosed with autism counting a child's desirable behavior of interest (e.g., communication) in each session (Symon, 2005). The data gathered using the AB design structure (4, 4, 5, 3, and 7 positive communications during baseline and 7, 8, 9, 7, and 9 during treatment phase) can be represented graphically as shown on Figure 1. In following section, the original and the proposed procedures are applied to the data set presented in order to illustrate their calculus.

INSERT FIGURE 1 ABOUT HERE

Percent of nonoverlapping data:

1) Identify the highest measurement in phase A. In the example it is 7 positive communications corresponding to baseline day 5.
2) Calculate the number of phase B data points that exceed the value identified in the previous step. The measurements corresponding to days 7, 8, and 10 are greater than 7, so there are 3 values exceeding phase A's highest value.
3) Divide the value obtained in step 2 by the number of observations in phase B. The number of phase B observations is 5 and the result of the division is 3/5 = .6.
4) Multiply the value obtained in step 3 by 100 in order to convert the proportion into a percentage. The percentage obtained for the example is .6 · 100 = 60%.

Percent of nonoverlapping corrected data:

1) Difference the phase A data points and obtain the differenced series with length n_A-1. In the example the differenced series has the following 5−1 = 4 data points: 0 (4−4), 1 (5−4), −2 (3−5), and 4 (7−3).
2) Compute the mean of the differenced series. The average of 0, 1, −2, and 4 is 0.75.
3) Compute the trend-correction factor for each data point: the mean of the differenced series multiplied by T_t. In the example the value of the correction factor are: .75 · 1, .75 · 2, …,.75 · 10.
4) Perform the data correction subtracting the corresponding correction factor from each original data point. After the correction phase A consists of 3.25 (4−.75 · 1), 2.5 (4−.75 · 2), 2.75 (5−.75 · 3), 0 (3−.75 · 4), and 3.25 (7−.75 · 5) and phase B the following data points: 2.5 (7−.75 · 6), 2.75 (8−.75 · 7), 3 (9−.75 · 8), .25 (7−.75 · 9), and 1.5 (9−.75 · 10).
5) Apply PND: None of the phase B data points is greater than the phase A highest value (3.25) and, therefore, PNCD = 0%.

Simulation

The specific steps that were implemented in the Fortran programs (one for each of the six series' length) were the following ones:

1) Systematic selection of each of the 19 values of φ_1 or θ_1.
2) Systematic selection of the (β_1, β_2, and β_3) parameters for data generation, leading to 8 different data patterns – autoregressive or moving average model with no effect or trend; trend; level change; slope change; trend

and level change; trend and slope change; combined level and slope change; trend and combined level and slope change.

3) 100,000 iterations of steps 4 through 15.

4) Generate the u_t term according to an exponential, a normal, or a uniform distribution, eliminating the first 50 random numbers using the next N ones.

5) Establish $\varepsilon_1 = u_1$.

6) Obtain the error term ε_t out of the random variable u_t using the AR(1) model $\varepsilon_t = \varphi_1 \cdot \varepsilon_{t-1} + u_t$ or the MA(1) model $\varepsilon_t = u_t - \theta_1 \cdot u_{t-1}$.

7) Obtain the time array $T_t = 1, 2, \ldots, N$.

8) Obtain the dummy treatment variable array D_t, where $D_t = 0$ for phase A and $D_t = 1$ for phase B.

9) Obtain the slope change array according to: $SC_t = [T_t - (n_A + 1)] \cdot D_t$.

10) Obtain the y_t array containing measurements (i.e., dependent variable): $y_t = \beta_0 + \beta_1 \cdot T_t + \beta_2 \cdot D_t + \beta_3 \cdot SC_t + \varepsilon_t$.

11) Calculate PND on the original data (i.e., the y_t array).

12) Correct data according to the procedure proposed.

13) Calculate PNCD on correct data.

14) Average the obtained percentages from the 100,000 replications of each experimental condition.

For data generation NAG libraries *nag_rand_neg_exp*, *nag_rand_normal*, and *nag_rand_uniform* were used. In order to guarantee suitable simulated data, the 50 values previous to each simulated data series were eliminated in order to reduce artificial effects (Greenwood & Matyas, 1990) and to avoid dependence between successive data series (Huitema, McKean, & McKnight, 1999).

Results

When the data series represent solely random fluctuation (i.e., there is no trend, autocorrelation, or treatment effect), the percentages provided by PNCD are systematically larger than the ones provided by PND, as illustrated by Figure 2. This finding implies that PND may be a better filter for ineffective interventions in absence of trend and serial dependence. In the abovementioned conditions, higher effect size estimates were also obtained for PNCD in comparison to PND when treatment effects existed. However, if data are present trend, the PND estimates increase and may become superior to the PNCD estimates for both independent (Figure 2) and serially related (Figure 3) data series, as the within-figure comparisons show.

INSERT FIGURES 2 AND 3 ABOUT HERE

Trend effect

In order to quantify the distortion of effect size estimates produced by trend, the ratio between percentages with and without trend in data was computed.

Therefore, a ratio close to 1 would indicate that trend does not introduce distortion, whereas values greater than 1 imply overestimation of the magnitude of effect. In the experimental conditions with no treatment effect simulated (Table 1) ratios > 1 entail an increment in false alarms, which is the case for PND in contrast with PNCD which maintains approximately the same magnitude estimates in presence and in absence of trend. This finding is applicable to all series lengths and errors' distributions tested.

INSERT TABLE 1 ABOUT HERE

When there is treatment effect (slope change, level change or both), the presence of trend leads to overestimation of the effect size obtained through PND, as Table 2 shows. In contrast, the estimates provided by PNCD are not affected by the confounding variable.

INSERT TABLE 2 ABOUT HERE

The ratios presented in Tables 1 and 2 show that the PND estimates become more distorted by trend when the number of measurements N increases. PNCD seems to deal effectively with trend for both shorter and longer data series.

Autocorrelation effect

The distortion of effect size estimates produced by serial dependence was quantified by means of the ratio between percentages computed for autocorrelated and independent data. Once again ratios of 1 imply no distortion and values greater than 1 are indicative of elevated false alarm rates in absence of intervention effect. In the case of exponential errors, for both AR(1) and MA(1) models PNCD performs worse than PND when there is negative autocorrelation, only slightly better for positive serial dependence. In contrast, for the normal and uniform errors, PNCD outperforms PND. For these two error distributions and AR(1) processes (Table 3) with $\varphi_1 > 0$ the difference between PNCD and PND increases for longer data series, whereas for $\varphi_1 < 0$ PNCD performs better only for $N \leq 20$. For the MA(1) processes (Table 4) with negative values of θ_1 (i.e., positive autocorrelation) PNCD shows less distortion than PND, whereas for $\theta_1 > 0$ it outperforms PND only for $N \leq 15$, always referring to normal and uniform errors.

INSERT TABLES 3 AND 4 ABOUT HERE

Combined effect

In addition to the individual effects of each of this data features, their combined effect was studied following the same procedure for quantifying distortion. Table 5 shows that for AR(1) processes with trend, PNCD is much less affected by the confounding variables than PND, whose effect size estimate is quintupled in certain experimental conditions. For MA(1) processes (Table 6), the findings are similarly favorable for PNCD.

INSERT TABLES 5 AND 6 ABOUT HERE

Discrimination between data patterns

In general the desirable characteristics of an effect size procedure are to be sensitive to intervention effects and not to be affected, for instance, by trend or serial dependence. Hence, an optimal performance (illustrated by Figure 4) would imply: a) low effect size estimates in absence of treatment effect; b) low effect size estimates when there is only general trend; c) higher estimates when there are actual changes in the response rate due to intervention.

INSERT FIGURE 4 ABOUT HERE

Comparing this ideal discrimination to the estimates obtain by means of PND and PNCD, it can be seen that there is a greater resemblance in the case of the latter procedure. That is, a combined effect (both change in level and in slope) yields a greater effect size estimate than an individual effect and the percentage obtained in absence of intervention effect is even lower. Additionally, trend does not shift estimates up as is the case for PND, which detects trend as an intervention effect. Figure 5 illustrates these findings for the shortest series length studied.

INSERT FIGURE 5 ABOUT HERE

Discussion

The present investigation proposes a data correction step to be introduced prior to applying the percent of nonoverlapping data as a technique for quantifying treatment effectiveness. The modified procedure is compared with the original in the context of data sets generated with known attributes such as trend, autocorrelation and treatment effect. For applied researchers, the results obtained suggest that PNCD is an effective method to deal with trend and can, therefore, be used in situations when pre-intervention measurements are not pure random fluctuation. Unstable baselines have been regarded as undesirable, but they can be common in applied settings where the introduction of the treatment is subjected to factors that cannot always be controlled by the practitioners. Although a professional might be reluctant to initiate the intervention when there is trend in data, treatment administration may be imposed by institutional time schedules, client's availability, etc. In such case, some kind of statistical control is advisable (Kazdin, 1978) and it can be achieved by means of the procedure proposed here. Apart from behavioral data with baseline trends, another potential context for application of PNCD are studies in which the data points are not sufficiently spaced in time and can present a sequential relation. PNCD ought to be preferred to PND in these cases, due to the fact that autocorrelation is more problematic for latter.

Whenever the behavioral measurements are not serially dependent and do not present trend, PND may be a better option than PNCD, since it produces lower magnitude of effect estimates. This difference in the estimates implies that in the abovementioned cases PND is less likely to label an intervention as effective when it is not. It has already been discussed that different effect size procedures may lead to different conclusions about the degree of treatment effectiveness for the same data set (McGrath & Meyer, 2006; Parker et al., 2005). In the particular case of PND and PNCD, the difference in estimates implies that the interpretation benchmarks proposed by Scruggs and Mastropieri (1998) cannot be applied directly to PNCD. On the other hand, there is evidence that PND is a conservative as compared to other procedures for estimating magnitude of effect (Jenson et al., 2007). Therefore, the effect size estimates provided by PNCD may resemble more the ones obtained by other models.

From a methodological perspective, PNCD can be regarded as an attempt to improve a procedure that is attractive to applied psychologists and is frequently employed by them. The aim is not only to achieve a better performance but also to maintain the simplicity of the technique. Therefore, we consider that the modifications balancing statistical properties improvements and low levels of calculus/interpretative complexity have to be encouraged. Furthermore, the present study follows the practice of offering data analysis programs for single-case designs in freeware like R (e.g., Bulté & Onghena, 2008); a practice we deem ought to be promoted.

The current investigation only focused on AB designs, although the results are potentially applicable to multiple-baseline designs (Busse, Kratochwill, & Elliott, 1995). The data sets used in the present study were constructed using permanent linear trend, constant variance and constant autocorrelation throughout the whole series. This data assumptions are common to simulation studies on N = 1 designs (e.g., Huitema & McKean, 2007a; 2007b; Matyas & Greenwood, 1990; Brossart et al., 2006; Parker & Brossart, 2003). Thus, future studies may explore the performance of PNCD for ABAB designs with curvilinear trends computing the percentage for each change in the condition as suggested by Kromrey and Foster-Johnson (1996). Additionally, comparative studies such as the present one which center on finding the technique that performs *better* need to be complemented by precision studies in order to identify techniques that perform *well*, that is, yield accurate estimates of the effect sizes simulated.

References

Allison, D. B., & Gorman, B. S. (1994). "Make things as simple as possible, but no simpler". A rejoinder to Scruggs and Mastropieri. *Behaviour Research and Therapy, 32*, 885-890.

Beretvas, S. N., & Chung, H. (2008). An evaluation of modified R^2-change effect size indices for single-subject experimental designs. *Evidence-Based Communication Assessment and Intervention, 2*, 120-128.

Bradley, J. V. (1977). A common situation conducive to bizarre distribution shapes. *American Statistician, 31*, 147-150.

Brossart, D. F., Parker, R. I., Olson, E. A., & Mahadevan, L. (2006). The relationship between visual analysis and five statistical analyses in a simple AB single-case research design. *Behavior Modification, 30*, 531-563.

Bulté, I., & Onghena, P. (2008). An R package for single-case randomization tests. *Behavior Research Methods, 40*, 467-478.

Busse, R. T., Kratochwill, T. R., & Elliott, S. N. (1995). Meta-analysis for single-case consultation outcomes: Applications to research and practice. *Journal of School Psychology, 33*, 269-285.

Cohen, J. (1990). Things I have learned (so far). *American Psychologist, 45*, 1304-1312.

Cohen, J. (1994). The earth is round (p < .05). *American Psychologist, 49*, 997-1003.

Crane, D. R. (1985). Single-case experimental designs in family therapy research: Limitations and considerations. *Family Process, 24*, 69-77.

Gedo, P. M. (2000). Single case studies in psychotherapy research. *Psychoanalytic Psychology, 16*, 274-280.

Greenwood, K. M., & Matyas, T. A. (1990). Problems with application of interrupted time series analysis for brief single-subject data. *Behavioral Assessment, 12*, 355-370.

Harrop, J. W., & Velicer, W. F. (1985). A comparison of alternative approaches to the analysis of interrupted time-series. *Multivariate Behavioral Research, 20*, 27-44.

Huitema, B. E., & McKean, J. W. (2000). Design specification issues in time-series intervention models. *Educational and Psychological Measurement, 60*, 38-58.

Huitema, B. E., & McKean, J. W. (2007a). An improved portmanteau test for autocorrelated errors in interrupted time-series regression models. *Behavior Research Methods, 39*, 343-349.

Huitema, B. E., & McKean, J. W. (2007b). Identifying autocorrelation generated by various error processes in interrupted time-series progression designs: A comparison of AR1 and portmanteau tests. *Educational and Psychological Measurement, 67*, 447-459.

Huitema, B. E., McKean, J. W., & McKnight, S. (1999). Autocorrelation effects on least-squares intervention analysis of short time series. *Educational and Psychological Measurement, 59*, 767-786.

Jenson, W. R., Clark, E., Kircher, J. C., & Kristjansson, S. D. (2007). Statistical reform: Evidence-based practice, meta-analyses, and single subject designs. *Psychology in the Schools, 44*, 483-493.

Kazdin, A. (1978). Methodological and interpretive problems of single-case experimental designs. *Journal of Consulting and Clinical Psychology, 46*, 629-642.

Kirk, R. E. (1996). Practical significance: A concept whose time has come. *Educational and Psychological Measurement*, *56*, 746-759.

Kratochwill, T. R., & Brody, G. H. (1978). Single subject designs: A perspective on the controversy over employing statistical inference and implications for research and training in behavior modification. *Behavior Modification, 2*, 291-307.

Kromrey, J. D., & Foster-Johnson, L. (1996). Determining the efficacy of intervention: The use of effect sizes for data analysis in single-subject research. *The Journal of Experimental Education*, *65*, 73-93.

Ma, H. H. (2006). An alternative method for quantitative synthesis of single-subject research: Percentage of data points exceeding the median. *Behavior Modification*, *30,* 598-617.

Manolov, R., & Solanas, A. (2008). Comparing N = 1 effect size indices in presence of autocorrelation. *Behavior Modification, 32*, 860-875.

Manolov, R., Solanas, A., & Leiva, D. (in press). Comparing "visual" effect size indices for single-case designs. *Methodology - European Journal of Research Methods for the Behavioral and Social Sciences.*

Matyas, T. A., & Greenwood, K. M. (1990). Visual analysis for single-case time series: effects of variability, serial dependence, and magnitude of intervention effects. *Journal of Applied Behavior Analysis, 23*, 341-351.

McCleary, R., & Hay, R. A., Jr. (1980). *Applied time series analysis for the social sciences.* Beverly Hills: Sage.

McGrath, R. E., & Meyer, G. J. (2006). When effect size disagree: The case of *r* and *d*. *Psychological Methods, 11*, 386-401.

Micceri, T. (1989). The unicorn, the normal curve, and other improbable creatures. *Psychological Bulletin, 105*, 156-166.

Parker, R. I., & Brossart, D. F. (2003). Evaluating single-case research data: A comparison of seven statistical methods. *Behavior Therapy, 34*, 189-211.

Parker, R. I., Brossart, D. F., Vannest, K. J., Long, J. R., Garcia De-Alba, R., Baugh, F. G., & Sullivan, J. R. (2005). Effect sizes in single case research: How large is large? *School Psychology Review, 34,* 116-132.

Parker, R. I., Cryer, J., & Byrns, G. (2006). Controlling baseline trend in single-case research. *School Psychology Quarterly, 21*, 418-443.

Parker, R. I., Hagan-Burke, S., & Vannest, K. (2007). Percentage of all non-overlapping data: An alternative to PND. *Journal of Special Education, 40*, 194-204.

Rosnow, R. L., & Rosenthal, R. (1989). Statistical procedures and the justification of knowledge in psychological science. *American Psychologist, 44,* 1276-1284.

Schlosser, R. W., Lee, D. L., & Wendt, O. (2008). Application of the percentage of non-overlapping data (PND) in systematic reviews and meta-analyses: A systematic review of reporting characteristics. *Evidence-Based Communication Assessment and Intervention, 2*, 163-187.

Schlosser, R. W., & Sigafoos, J. (2008). Meta-analysis of single-subject designs: Why now? *Evidence-Based Communication Assessment and Intervention, 2*, 117-119.

Schneider, N., Godlstein, H., & Parker, R. (2008). Social skills interventions for children with autism: A meta-analytic application of percentage of all non-overlapping data (PAND). *Evidence-Based Communication Assessment and Intervention, 2*, 152-162.

Scruggs, T. E., & Mastropieri, M. A. (1998). Summarizing single-subject research: Issues and applications. *Behavior Modification*, *22*, 221-242.

Scruggs, T. E., Mastropieri, M. A., & Casto, G. (1987). The quantitative synthesis of single-subject research: Methodology and validation. *Remedial and Special Education*, 8, 24-33.

Shadish, W. R., Rindskopf, D. M., & Hedges, L. V. (2008). The state of the science in the meta-analysis of single-case experimental designs. *Evidence-Based Communication Assessment and Intervention, 2*, 188-196.

Symon, J. B. (2005). Expanding interventions for children with autism: Parents as trainers. *Journal of Positive Behavior Interventions, 7*, 159-173.

Tervo, R. C., Estrem, T. L., Bryson-Brockman, W., & Symons, F. J. (2003). Single-case experimental designs: Application in developmental-behavioral pediatrics. *Developmental and Behavioral Pediatrics, 24*, 438-448.

Wilkinson, L., & The Task Force on Statistical Inference. (1999). Statistical methods in psychology journals: Guidelines and explanations. *American Psychologist*, *54,* 694-704.

Appendix I

R code computing PND and PNCD as output. The input required from the user is: 1) the data for phase A in the expression **phaseA <- c(1:10),** replacing "**1:10**" with the measurements obtained separated by commas; and 2) the data for phase B placed instead of "**11:20**" in the expression **phaseB <- c(11:20)**. After introducing the behavioral measurements, the text is copied and pasted into the R console and the estimates are printed out.

```
# Data input
phaseA <- c(1:10)
phaseB <- c(11:20)
n_a <- length(phaseA)
n_b <- length(phaseB)

# Data correction: phase A
phaseAdiff <- c(1:(n_a-1))
for (iter1 in 1:(n_a-1))
phaseAdiff[iter1] <- phaseA[iter1+1] - phaseA[iter1]
phaseAcorr <- c(1:n_a)
for (iter2 in 1:n_a)
phaseAcorr[iter2] <- phaseA[iter2] - mean(phaseAdiff)*iter2

# Data correction: phase B
phaseBcorr <- c(1:n_b)
for (iter3 in 1:n_b)
phaseBcorr[iter3] <- phaseB[iter3] - mean(phaseAdiff)*(iter3+n_a)

# PND on corrected data
countcorr <- 0
for (iter4 in 1:n_b)
if (phaseBcorr[iter4] > max(phaseAcorr)) countcorr <- countcorr+1
pndcorr <- (countcorr/n_b)*100
print ("The percentage of nonoverlapping corrected data is"); print(pndcorr)

# PND on original data
count <- 0
for (iter5 in 1:n_b)
if (phaseB[iter5] > max(phaseA)) count <- count+1
pnd <- (count/n_b)*100
print ("The percent of nonoverlapping data is"); print(pnd)
```

Appendix II

R code computing PND and PNCD as output used as described in Appendix I. Useful when the objective of the behavior of interest is an undesirable one and the treatment pretends to eliminate or reduce it.

```
# Data input
phaseA <- c(1:10)
phaseB <- c(11:20)
n_a <- length(phaseA)
n_b <- length(phaseB)

# Data correction: phase A
phaseAdiff <- c(1:(n_a-1))
for (iter1 in 1:(n_a-1))
phaseAdiff[iter1] <- phaseA[iter1+1] - phaseA[iter1]
phaseAcorr <- c(1:n_a)
for (iter2 in 1:n_a)
phaseAcorr[iter2] <- phaseA[iter2] - mean(phaseAdiff)*iter2

# Data correction: phase B
phaseBcorr <- c(1:n_b)
for (iter3 in 1:n_b)
phaseBcorr[iter3] <- phaseB[iter3] - mean(phaseAdiff)*(iter3+n_a)

# PND on corrected data
countcorr <- 0
for (iter4 in 1:n_b)
if (phaseBcorr[iter4] < min(phaseAcorr)) countcorr <- countcorr+1
pndcorr <- (countcorr/n_b)*100
print ("The percentage of nonoverlapping corrected data is"); print(pndcorr)

# PND on original data
count <- 0
for (iter5 in 1:n_b)
if (phaseB[iter5] < min (phaseA)) count <- count+1
pnd <- (count/n_b)*100
print ("The percent of nonoverlapping data is"); print(pnd)
```

Tables

Table 1. Distortion due to trend in independent data series – the values represent the ratio between presence of trend / absence of trend in experimental conditions without treatment effect.

Phase length		Ratio trend / random fluctuations	
n_A	n_B		
exponential		PND	PNCD
5	5	1.336	.996
5	10	1.576	1.002
7	8	1.570	1.003
10	10	1.807	.999
15	15	2.431	1.000
20	20	3.293	.995
normal		PND	PNCD
5	5	1.429	.998
5	10	1.674	1.000
7	8	1.772	.997
10	10	2.279	1.000
15	15	3.601	1.003
20	20	5.511	1.005
uniform		PND	PNCD
5	5	1.517	1.002
5	10	1.747	.997
7	8	2.003	.995
10	10	2.761	1.005
15	15	4.590	.991
20	20	6.952	.993

Table 2. Distortion due to trend in independent data series – the values represent the ratio between presence of trend / absence of trend in experimental conditions with single or combined treatment effect.

Phase length		Ratio trend & level / level change only		Ratio trend & slope / slope change only		Ratio trend & both effects / both effects	
n_A	n_B						
exponential		PND	PNCD	PND	PNCD	PND	PNCD
5	5	1.338	.999	1.340	.996	1.301	1.004
5	10	1.547	1.003	1.380	.994	1.298	1.002
7	8	1.544	.991	1.507	1.003	1.433	1.002
10	10	1.803	.992	1.723	1.005	1.605	1.000
15	15	2.433	1.001	1.962	.999	1.782	.998
20	20	3.240	.998	1.992	.998	1.808	1.007
normal		PND	PNCD	PND	PNCD	PND	PNCD
5	5	1.353	1.002	1.348	1.002	1.287	.995
5	10	1.546	1.010	1.385	.999	1.301	1.002
7	8	1.627	1.005	1.523	1.005	1.413	.995
10	10	2.016	.999	1.703	1.003	1.547	.998
15	15	2.985	1.004	1.779	.996	1.601	1.003
20	20	4.220	.998	1.681	1.003	1.525	.997
uniform		PND	PNCD	PND	PNCD	PND	PNCD
5	5	1.325	1.001	1.339	.998	1.246	.998
5	10	1.506	.996	1.354	.997	1.276	.997
7	8	1.596	.995	1.456	.996	1.350	1.005
10	10	1.909	1.008	1.562	.997	1.432	.998
15	15	2.530	.996	1.611	1.003	1.473	1.001
20	20	3.119	.989	1.505	1.000	1.374	.994

Table 3. Distortion due to an AR(1) process – the values represent the ratio between serially dependent data and independent series with no trend or intervention effect.

Phase length		Ratio φ_1=−.3 / random fluctuations		Ratio φ_1=.3 / random fluctuations		Ratio φ_1=.6 / random fluctuations	
n_A	n_B						
exponential		PND	PNCD	PND	PNCD	PND	PNCD
5	5	.941	.926	1.135	1.121	1.302	1.250
5	10	.948	.943	1.169	1.110	1.422	1.234
7	8	.955	.943	1.167	1.147	1.482	1.365
10	10	.958	.940	1.164	1.157	1.559	1.455
15	15	.956	.950	1.138	1.157	1.591	1.511
20	20	.981	.953	1.141	1.152	1.614	1.545
normal		PND	PNCD	PND	PNCD	PND	PNCD
5	5	.933	.953	1.158	1.065	1.379	1.121
5	10	.933	.968	1.167	1.042	1.441	1.093
7	8	.946	.955	1.171	1.067	1.503	1.173
10	10	.944	.953	1.174	1.069	1.579	1.207
15	15	.965	.954	1.178	1.075	1.637	1.221
20	20	.983	.957	1.168	1.065	1.634	1.212
uniform		PND	PNCD	PND	PNCD	PND	PNCD
5	5	.929	.961	1.158	1.060	1.378	1.114
5	10	.929	.967	1.166	1.032	1.428	1.079
7	8	.936	.954	1.185	1.050	1.497	1.130
10	10	.932	.946	1.195	1.050	1.561	1.146
15	15	.949	.920	1.189	1.006	1.602	1.088
20	20	.971	.912	1.188	.994	1.619	1.035

Table 4. Distortion due to an MA(1) process – the values represent the ratio between nonnull and null θ_1 parameters in series with no trend or intervention effect.

Phase length		Ratio θ_1=−.5 / random fluctuations		Ratio θ_1=.5 / random fluctuations	
n_A	n_B				
exponential		PND	PNCD	PND	PNCD
5	5	.913	.887	1.232	1.177
5	10	.903	.914	1.242	1.151
7	8	.912	.886	1.248	1.190
10	10	.915	.882	1.257	1.196
15	15	.925	.910	1.240	1.206
20	20	.947	.920	1.235	1.196
normal		PND	PNCD	PND	PNCD
5	5	1.203	1.077	.887	.927
5	10	1.217	1.063	.880	.947
7	8	1.226	1.080	.901	.919
10	10	1.221	1.077	.910	.902
15	15	1.207	1.067	.931	.905
20	20	1.197	1.058	.945	.899
uniform		PND	PNCD	PND	PNCD
5	5	1.194	1.066	.882	.938
5	10	1.200	1.046	.868	.951
7	8	1.207	1.047	.881	.925
10	10	1.205	1.038	.906	.894
15	15	1.218	.990	.929	.860
20	20	1.179	.947	.940	.842

Table 5. Distortion due to combined presence of trend and an AR(1) process – the values represent the ratio between serially dependent data with trend and independent series with no trend.

Phase length		Ratio trend & φ_1=−.3 / random fluctuations		Ratio trend & φ_1=.3 / random fluctuations		Ratio trend & φ_1=.6 / random fluctuations	
n_A	n_B						
exponential		PND	PNCD	PND	PNCD	PND	PNCD
5	5	1.267	.929	1.518	1.131	1.671	1.239
5	10	1.484	.942	1.785	1.116	2.007	1.240
7	8	1.489	.941	1.793	1.140	2.012	1.284
10	10	1.734	.939	2.074	1.164	2.566	1.459
15	15	2.336	.951	2.734	1.150	3.368	1.524
20	20	3.211	.960	3.679	1.141	4.369	1.520
normal		PND	PNCD	PND	PNCD	PND	PNCD
5	5	1.325	.953	1.581	1.056	1.783	1.126
5	10	1.565	.969	1.839	1.036	2.043	1.089
7	8	1.663	.950	1.961	1.068	2.223	1.170
10	10	2.130	.944	2.457	1.068	2.763	1.210
15	15	3.359	.944	3.768	1.068	3.961	1.200
20	20	5.100	.955	5.594	1.075	5.585	1.215
uniform		PND	PNCD	PND	PNCD	PND	PNCD
5	5	1.370	.959	1.600	1.054	1.765	1.115
5	10	1.587	.968	1.826	1.032	1.980	1.076
7	8	1.795	.952	2.056	1.049	2.205	1.133
10	10	2.415	.945	2.711	1.054	2.758	1.139
15	15	4.027	.917	4.360	1.015	4.033	1.092
20	20	6.132	.904	6.590	.995	5.766	1.037

Table 6. Distortion due to combined presence of trend and an MA(1) process – the values represent the ratio between moving average data with trend and data series with $\theta_1 = 0$ and no trend.

Phase length n_A	n_B	Ratio trend & θ_1=−.5 / random fluctuations PND	PNCD	Ratio trend & θ_1=.5 / random fluctuations PND	PNCD
exponential		PND	PNCD	PND	PNCD
5	5	1.217	.888	1.597	1.174
5	10	1.406	.910	1.837	1.153
7	8	1.418	.892	1.874	1.196
10	10	1.658	.877	2.153	1.184
15	15	2.268	.910	2.839	1.201
20	20	3.099	.920	3.727	1.182
normal		PND	PNCD	PND	PNCD
5	5	1.608	1.075	1.253	.924
5	10	1.850	1.062	1.462	.947
7	8	1.962	1.078	1.562	.919
10	10	2.412	1.072	1.970	.909
15	15	3.619	1.066	3.074	.900
20	20	5.231	1.052	4.601	.895
uniform		PND	PNCD	PND	PNCD
5	5	1.602	1.082	1.271	.933
5	10	1.831	1.054	1.485	.949
7	8	1.974	1.056	1.617	.924
10	10	2.539	1.039	2.150	.891
15	15	3.938	.992	3.503	.857
20	20	5.815	.949	5.281	.836

Figures

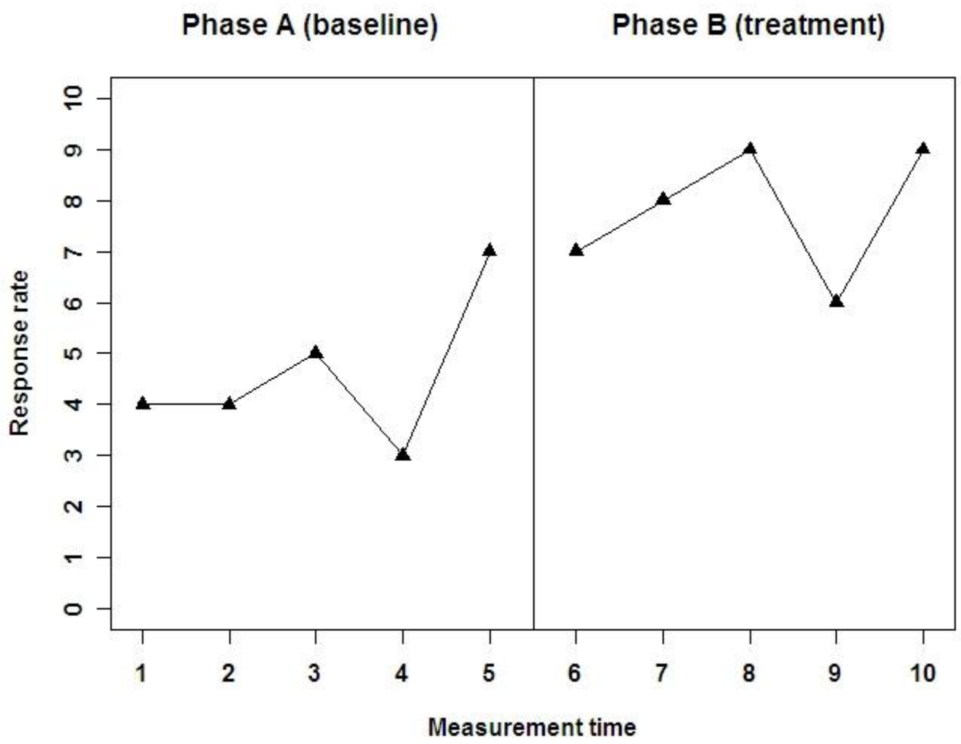

Figure 1. A fictitious example of an AB data series with $n_A = n_B = 5$.

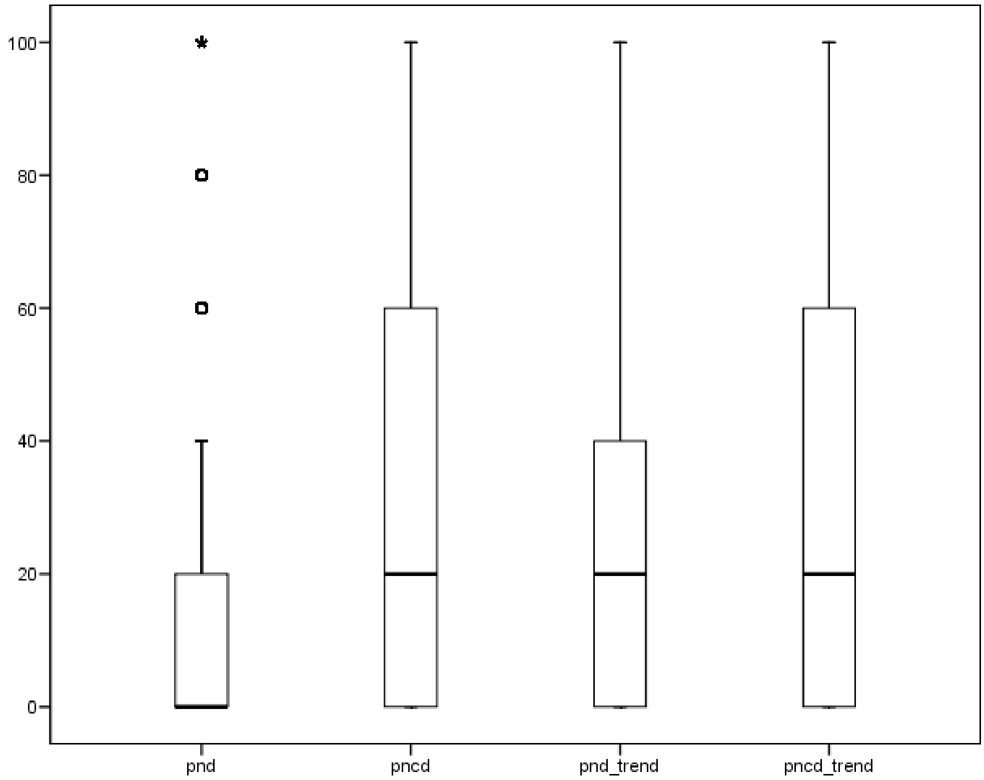

Figure 2. Distribution of the percentages provided by PND and PNCD in absence (the two box plots on the left) and presence of trend (the two box plots on the right). 100,000 samples of independent $n_A = n_B = 5$ data with no treatment effect simulated and normal error.

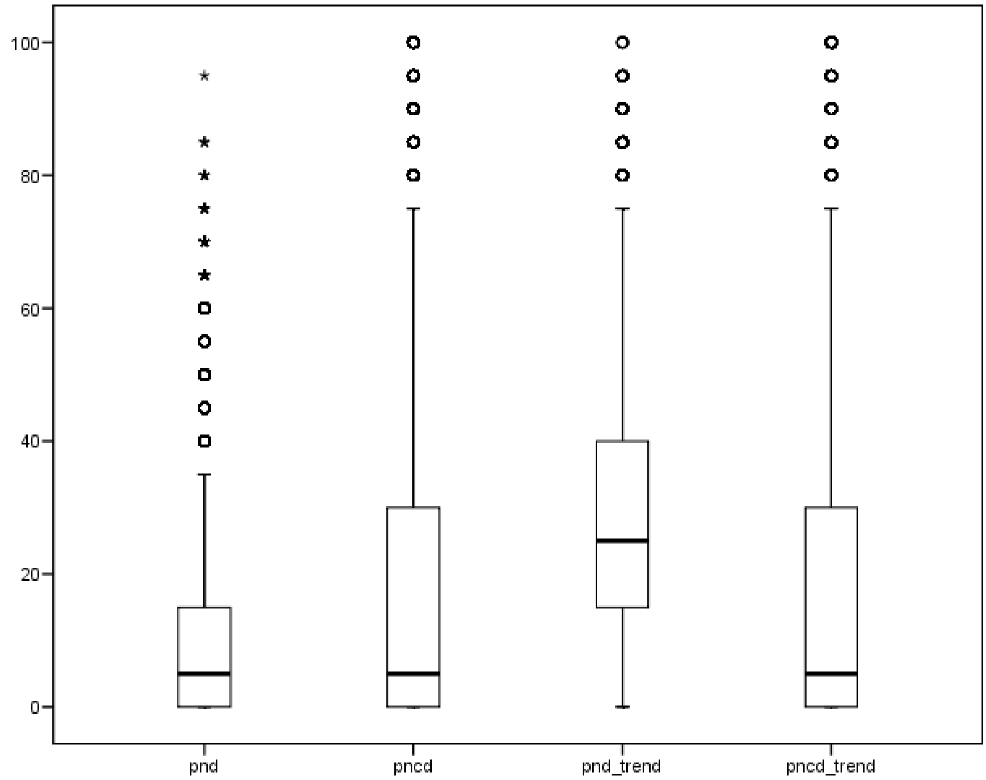

Figure 3. Distribution of the percentages provided by PND and PNCD in absence (the two box plots on the left) and presence of trend (the two box plots on the right). 100,000 samples of $n_A = n_B = 20$ data with level change simulated and uniform error in moving average processes with autocorrelation of .5.

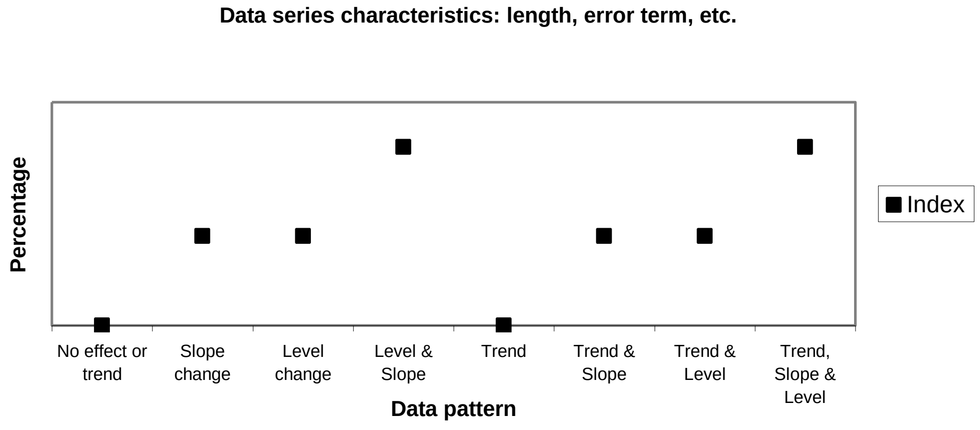

Figure 4. Ideal discrimination between data patterns.

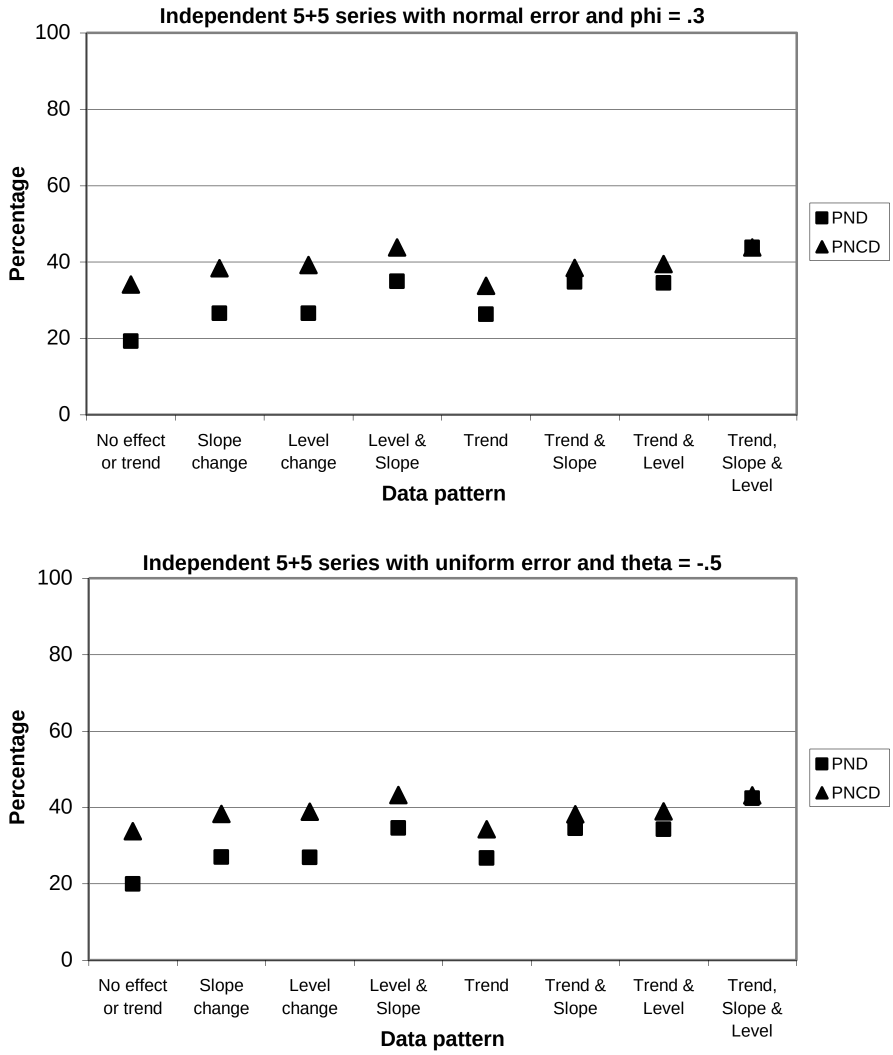

Figure 5. Discrimination between data patterns for both indices in different experimental conditions. Upper panel: N = 10 series generated from an AR process with normal error and φ_1=.3. Lower panel: N = 10 series generated from an MA process with uniform error and θ_1=−.5.