DIGITALCOMMONS
—@WAYNESTATE—

Journal of Modern Applied Statistical Methods

11-1-2013

The Single-Case Data Analysis Package: Analysing Single-Case Experiments with R Software

Isis Bulté
KU Leuven, Belgium, isis.bulte@ppw.kuleuven.be

Patrick Onghena
KU Leuven, Belgium, patrick.onghena@ppw.kuleuven.be

Recommended Citation

Statistical Software Applications and Review:

The Single-Case Data Analysis Package: Analysing Single-Case Experiments with R Software

Isis Bulté
KU Leuven
Belgium

Patrick Onghena
KU Leuven
Belgium

The RcmdrPlugin.SCDA plug-in package is discussed. It integrates three R packages in the R commander interface: SCVA (for Single-Case Visual Analysis), SCRT (for Single-Case Randomization Tests), and SCMA (for Single-Case Meta-Analysis). This way the plug-in package covers three important steps in the analysis of single-case data.

Keywords: Single-case studies, data analysis, software, R package, R commander plugin, GUI

Introduction

To investigate research questions in educational, behavioral, and medical science, single-case experiments are very well suited. To bring these experiments to the attention of (applied) researchers a software package is suggested to analyze data resulting from single-case experiments.

Single-case designs are increasingly popular in educational, behavioral, and medical research (Hammond & Gast, 2010; Matson, Turygin, Beighley, & Matson, 2012; Shadish & Sullivan, 2011; Swaminathan & Rogers, 2007). Bliss, Skinner, Hatau, and Carroll (2008), for example, classified all articles published in four school psychology journals (*School Psychology Quarterly, School Psychology Review, Journal of School Psychology,* and *Psychology in the Schools*) between 2000 and 2005 and found that, with the exception of 2004,

Dr. Bulté is a Researcher in the Faculty of Psychology and Educational Sciences, Methodology of Educational Sciences. Email her at: isis.bulte@ppw.kuleuven.be. Dr. Onghena is a full professor in the Faculty of Psychology and Educational Sciences, Methodology of Educational Sciences. Email him at patrick.onghena@ppw.kuleuven.be.

single-subject designs were more prevalent than group designs. In this time frame, 55% of the causal-experimental studies in these journals utilized single-subject designs. This may seem odd given the strong emphasis on large N designs in most statistical and methodological courses and handbooks for the educational, behavioral, and medical sciences.

A possible explanation for the findings of Bliss et al. (2008) lies in the fact that single-case designs can provide a viable alternative or supplement to group designs to answer causal questions. This can be said about educational, behavioral, and medical research in general, but the published applications of single-case designs are certainly not that predominantly present in all subareas. In many subareas of educational, behavioral, and medical research single-case designs however there is a huge potential for single-case designs to complement the standard group designs because of the necessity of pilot data in early stages of larger group studies, because of the relevance of research concerning rare types of participants (e.g., patients with a very specific neuropsychological disorder due to brain injury), or the examination of idiographic questions like 'does this intervention (e.g., restructuring) work for this particular organization?'; and, of course, when research funds are scarce and it is not possible to obtain enough participants for large-scale group studies (Barlow et al., 2009; Edgington & Onghena, 2007; Franklin, Allison, & Gorman, 1997; Kazdin, 2011). Because of the close link of single-case evaluations to individual care, single-case designs are also ideally suited to bridge the scientist-practitioner gap (Barlow, Hayes, & Nelson, 1984; Bliss et al., 2008).

Unfortunately, most of the commonly used statistical software packages, like SPSS and SAS, do not present readily available procedures or options for designing single-case experiments and analyzing single-case data. To fill this gap, an R package for designing single-case experiments and analyzing single-case data is presented.

R was chosen as the computational environment, because it is open source software, running on a variety of UNIX platforms, as well as on Windows and MacOS (Hornik, 2012). R has excellent graphical possibilities, but is also a very powerful and flexible statistical environment, which facilitates the combination of visual and statistical data analysis (Kelley, 2007). However, because of the use of a standard command line interface, R is not very user-friendly. Especially for practitioners who never engaged in any basic programming, the threshold to start working with R can be too high.

Fortunately, Fox (2005) created a window-based graphical user interface (GUI) to R, called the "R Commander" with menus and dialog boxes (very

similar to, e.g., SPSS), which are far more familiar to most people (see Figure 1). Working with the R commander is hands-down: by selecting the menus, submenus open which lead to dialog boxes. Each dialog box contains a 'help' button, which refers the user to a help page with more information. By making selections in the dialog boxes, R commands are generated and executed.
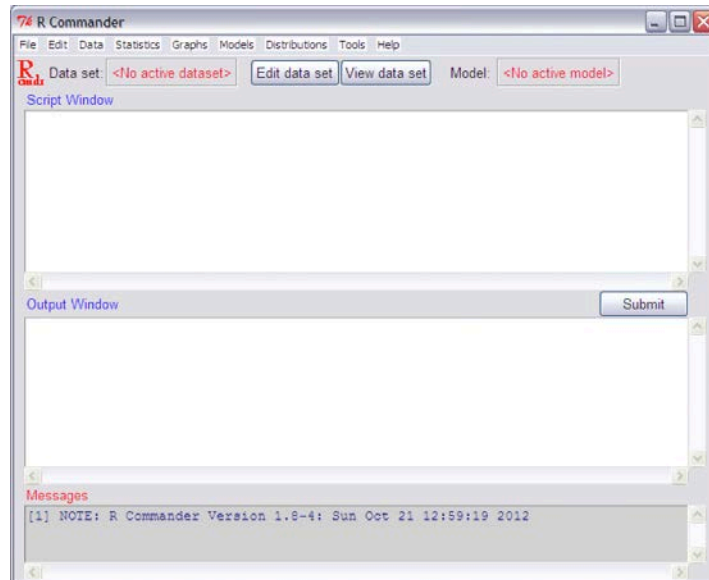

Figure 1. The R commander interface upon starting up.

An advantage of the R commander is that these R commands are not just executed 'behind the scene'. They also appear in the *script window,* so R users can always see the code and adapt it if necessary. This visibility of commands can also be a useful first step in learning R, by gently getting acquainted with the R language. In this script window, also other commands can be typed, or previous commands can be rerun. In the *output window*, the given command appears together with the output, and in the *messages window*, error messages, warnings, and notes appear (Fox, 2005).

An additional advantage of the R commander is that users can add their own menus and dialog boxes. This extensibility became even more practical with the possibility of writing plug-in packages (Fox, 2007). Besides the already mentioned advantages of R and the R Commander, a huge advantage is that all functionalities that are already available in the standard R Commander and in

other plug-in packages can be used. Instead of developing a stand-alone GUI, therefore a plug-in for the R Commander was created. The SCDA (Single Case Data Analysis) plug-in created is a GUI for three R packages presented: SCVA (Single Case Visual Analysis; Bulté & Onghena, 2012), SCRT (Single Case Randomization Tests; Bulté & Onghena, 2008, 2009) and SCMA (Single Case Meta-Analysis; Bulté et al., submitted).

The RcmdrPlugin.SCDA Package: An Illustration

When planning an experiment, first the study design should be carefully chosen based on, among other things, the research question. Then data can be collected according to the selected design. Single-case data analysis, just like any other analysis of empirical data, best starts with a visual exploration of the data. Unless the visualization is that convincing that the effect of the intervention is very obvious (e.g., a dramatic shift in level without any variability or trend), statistical data analysis might be a useful supplementary technique. And often it is not only useful to know whether an intervention had a statistically significant effect, but also what the size of the effect was. The SCDA plug-in for the R commander provides R functions for each of those steps of research design and data analysis. It adds one menu item to the R commander ("SCDA"), which contains three submenus (see Figure 2). Each of the submenus leads to several menu items.
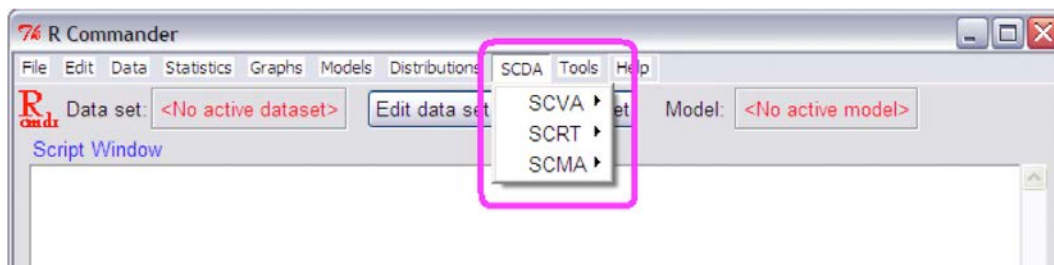

Figure 2. The R commander with the SCDA plugin.

1)   SCVA (Single-Case Visual Analysis)
- Graphical display
- Plot measure of central tendency
- Plot estimate of variability

- Plot estimate of trend

2) SCRT (Single-Case Randomization Tests)

- Design your experiment
  - o Number of possible assignments
  - o Display all possible assignments
  - o Choose 1 possible assignment

- Analyze your data
  - o Observed test statistic
  - o Randomization distribution
  - o P-value

3) SCMA (Single-Case Meta-Analysis)

- Calculate effect size
- Combine *p*-values

What follows will illustrate the functionality of the package with an example from ter Kuile et al. (2009). A more detailed manual can be found in the Appendix.

Ter Kuile et al. (2009) investigated the effectiveness of therapist-aided exposure for lifelong vaginismus. Lifelong (or primary) vaginismus occurs when a woman has never been able to have sexual penetrative intercourse during her whole life. From a cognitive-behavioral perspective, fear and avoidance behavior are connected to vaginismus. Therefore it was hypothesized that exposure to the feared stimuli (i.e., penetration) would increase the ability to have sexual intercourse. The exposure therapy was aided by a trained female therapist and consisted of vaginal penetration exercises (performed by the woman herself) at the hospital, together with several specific homework assignments in which the partner was involved.

Research Design and Data Collection

Ter Kuile et al. (2009) used a replicated single-case AB-phase design to test whether the therapy would lead to an increase in sexual intercourse. In an AB-phase design all baseline measurements (A) precede all treatment measurements

(B). The idea behind this design is to be able to attribute a change in intercourse behavior after the exposure onset to the therapy. To control for time-related confounding variables, randomization was incorporated in the study design: the random aspect was the start of the exposure therapy. This random determination of therapy onset is of course not unlimited. To decide whether the therapy had any effect on the ability to have intercourse, baseline as well as treatment observations are needed. Therefore, in this illustration a constraint of a minimum of one week of diary recordings in each phase is guarded. With a total of 24 weeks of data recordings, this leads to 23 possible start points for the therapy (after week 1, after week 2, …, after week 23) (Figure 3: SCRT -> Design your experiment -> Number of possible assignments).
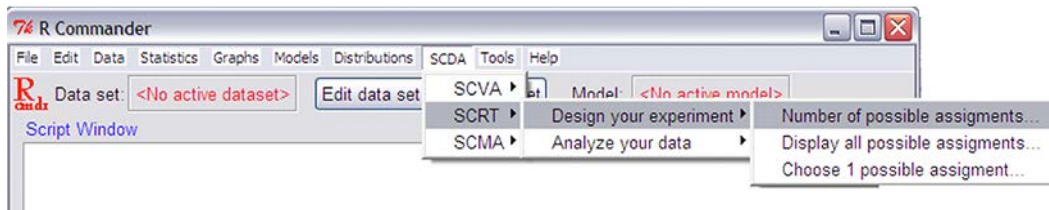

Figure 3. The RcmdrPlugin.SCDA menu, with the SCRT (Single-Case Randomization Tests) 'design your experiment' submenu.

By replicating this experiment over several participants, the strength of the findings is increased and more general statements can be made. Ten patients in sexology clinics, who suffered from primary vaginismus, participated in the study. They kept a daily diary in which they noted (amongst other things) whether they were able to have sexual intercourse with their partner that day. To analyze these data, they should be put into the R commander. This can be accomplished by reading in a created text file with the observations or by entering the observations directly as an active data set (see Appendix).

First the focus is on the diary data of the first patient (P1). She started recording her sexual intercourse attempts from the moment she was referred to the outpatient sexology clinic. The start of the exposure therapy was randomly determined after seven weeks of 'baseline' and she continued filling in the diary for seventeen more weeks.
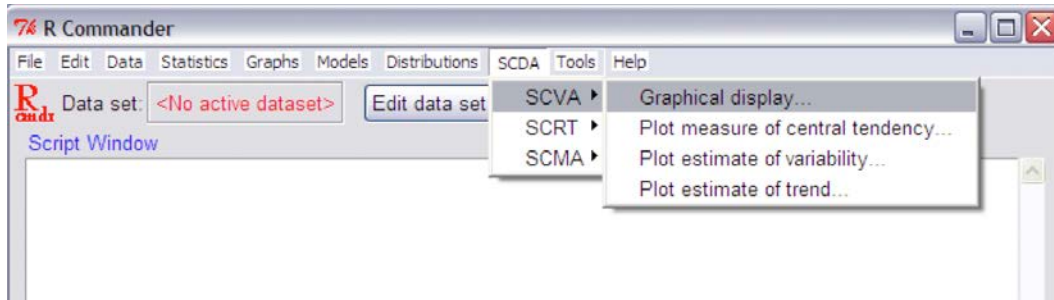
SCVA (Single-Case Visual Analysis)

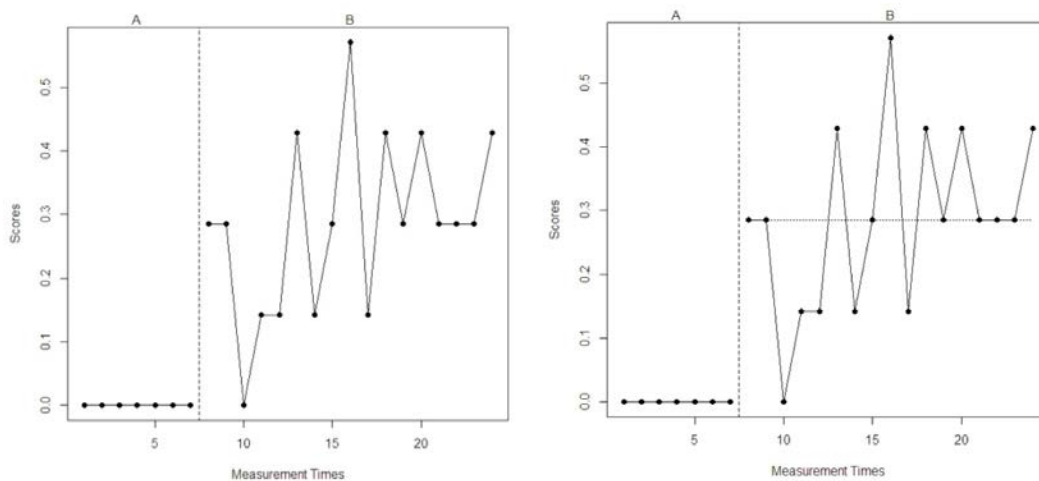**Figure 4.** The RcmdrPlugin.SCDA menu, with the SCVA (Single-Case Visual Analysis) submenu.

Figure 5. Graphical representation of the diary data of P1, averaged per week (SCDA -> SCVA -> Graphical display; see Figure 4). On the right, the shift in mean level is visualized (SCDA -> SCVA -> Plot measure of central tendency; see Figure 4).

A first impression can be obtained by making a graphical representation of her data (SCDA -> SCVA -> Graphical display; see Figure 4). The left panel of Figure 5 shows the average number of successful intercourse attempts per week for P1. The dotted vertical line indicates the start of the intervention phase. The introduction of the exposure therapy clearly made a difference: before therapy this woman had never experienced sexual intercourse with penetration. The shift in

456

mean level between the phases is illustrated in the right panel of Figure 5, where the phase means are plotted as a horizontal reference line on the raw data.
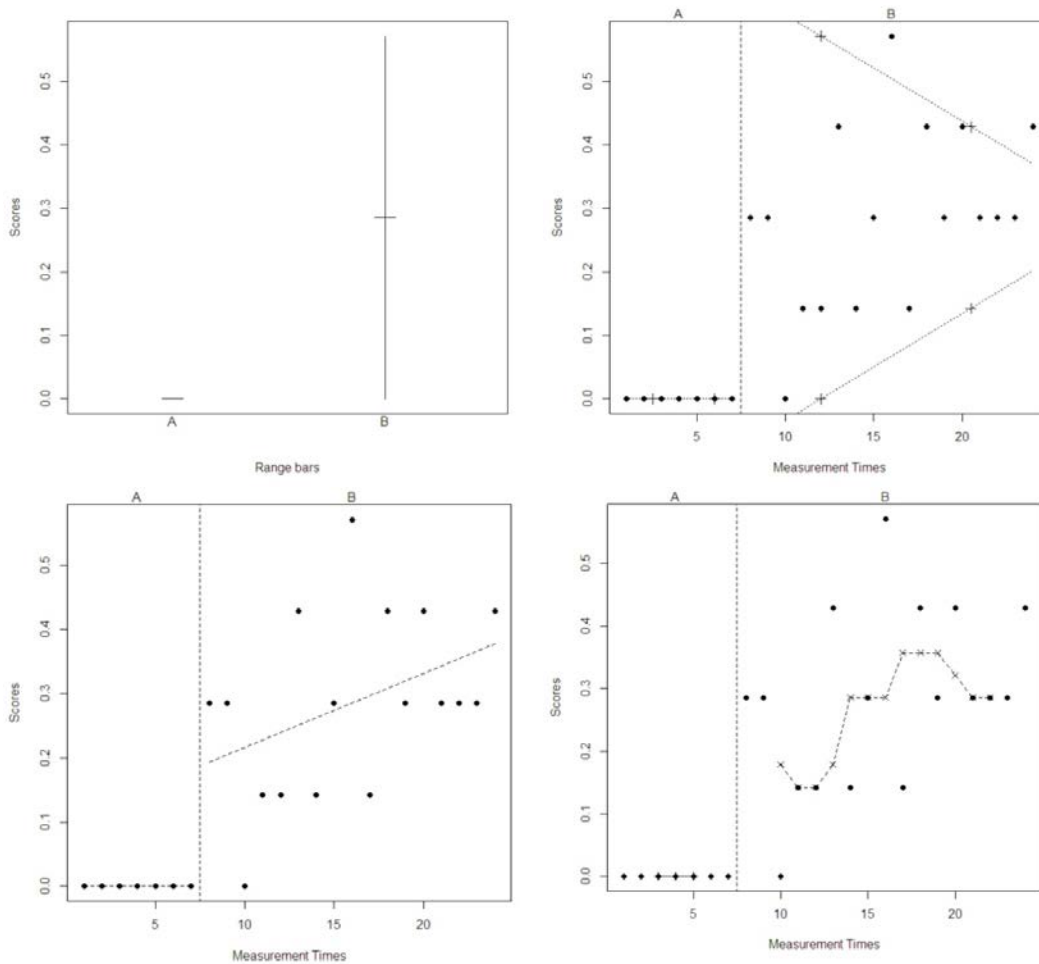

Figure 6. The upper panels illustrate the variation in the data (SCDA -> SCVA -> Plot estimate of variability; see Figure 4): range bars on the left and trended range lines on the right. Possible trends in the data are made visible in the lower panels (SCDA -> SCVA -> Plot estimate of trend; see Figure 4). On the left a linear function is drawn on the raw data by means of ordinary least-squares regression and on the right a nonlinear trend is visualized by displaying running medians of batch size four averaged by pairs.

The range bars in the upper left panel of Figure 6 illustrate the lack of variability in the baseline phase. In the treatment phase variation is shown by the vertical line connecting the minimum and the maximum value, with a small

horizontal bar marking the phase mean. The trended range lines depicted in the upper right panel of Figure 6 show that the variability decreases over time. There is also an upward trend in level noticeable from the linear function in the lower left panel of Figure 6. That this trend is not entirely linear can be seen from the nonlinear smoothed curve produced by calculating running medians of batch size four and averaging each successive pair.

SCRT (Single-Case Randomization Tests)

Visual analysis is an important first step when evaluating intervention effects. In addition, several statistical tests might be conducted to evaluate the statistical significance of the intervention effects. The SCDA GUI includes functions to conduct randomization tests: permutation tests based on random assignment, to test a null hypothesis about treatment effects in a randomized experiment (Onghena & Edgington, 2005). The alternative hypothesis in the illustration presented was that the exposure therapy will lead to an increase in successful sexual intercourse attempts. In other words, the mean of the treatment observations is expected to be higher than the mean of the baseline observations. Therefore, the difference between those means was chosen as test statistic.
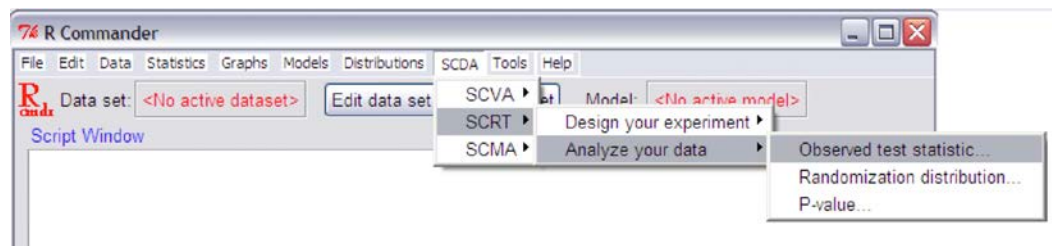
Figure 7. The RcmdrPlugin.SCDA menu, with the SCRT (Single-Case Randomization Tests) 'analyze your data' submenu.

Visual analysis of the data already indicated that the exposure therapy had a positive effect on the intercourse frequency of P1. This effect is also statistically significant, shown by the randomization test's *p*-value of .04 (Figure 7: SCRT -> Analyze your data -> P-value). More information on randomization tests can be found, for example, in Bulté and Onghena (2008).

SCMA (Single-Case Meta-Analysis)

The combination of visual analysis and statistical significance testing, however, does not tell the whole story. It is often also useful to know what the size of the effect is. How large an effect is, is expressed by means of effect size measures (Figure 8: SCDA -> SCMA -> Calculate effect size). The pooled standardized mean difference, which uses the pooled standard deviation of both phases, for P1 equals 2.35. The percentage of data in the treatment phase that is higher than the median of the baseline phase (=”PEM”) is 94%.
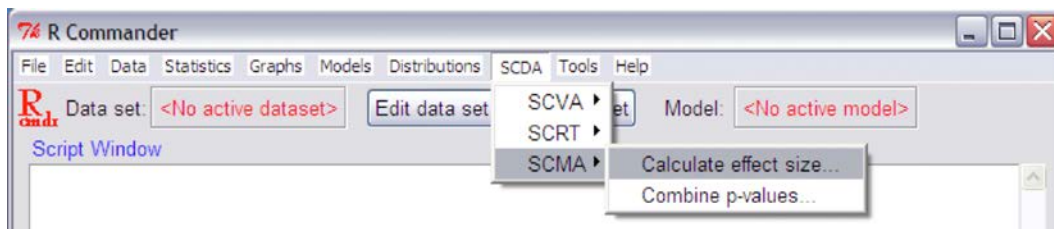

Figure 8. The RcmdrPlugin.SCDA menu, with the SCMA (Single-Case Meta-Analysis) submenus.

These effect sizes are not only used to show the magnitude of an effect, but they also help comparing results over different studies and they can be combined in meta-analyses to come to one general estimate of the effect size. Another meta-analytical procedure is statistically combining the *p*-values of several sequentially replicated single-case experiments, to determine whether a general significant result is obtained (Figure 8: SCDA -> SCMA -> Combine P-values). Ten women with primary vaginismus participated in the study of ter Kuile et al. (2009). When combining their results with the additive method for combining independent *p*-values, a general significant result of *p* = .0016 was obtained.

Conclusion

The use of the SCDA plugin package for the R commander was illustrated. A more elaborated overview of all functionalities and an explanation of the different parts in the dialogue windows is given in the Appendix.

The presented software package contains one technique for statistically analyzing the data resulting from single-case experiments. These randomization tests have several advantages over other techniques suggested in the literature.

Most importantly, the presence of serial dependency in the data will not invalidate the result of a randomization test. This is of considerable interest in the context of single-case experiments, because these data tend to have autocorrelated residuals, which can seriously bias the results of for example *t*-tests, by inflating the Type I error rate. Another advantage is that they are free from the assumption of random sampling. This random sampling assumption, on which the probability tables of parametric tests are built, is usually an unrealistic ideal and most experiments do not use randomly sampled subjects. Also, unlike parametric *t*-test, randomization tests are not based on assumptions about the homogeneity of variances. But they do have one requirement: the incorporation of random assignment not only enhances the internal validity of the study, as indicated before, but it also justifies the application of a randomization test that is based on the random assignment used in the experimental design of the study. This way it is possible to infer causal relations between the treatment and the observed changes in behavior. Because randomization tests acquire validity by mirroring the random assignment schedule used in the study, their extreme versatility permits researchers to make valid tests by devising a test that is suitable for the particular design used. They can be used with all sorts of data (continuous, discrete, ranks,...) and all kinds of data patterns (trends, outliers, skewed distributions, zero variance in baseline,...) (see e.g., Edgington, 1973, 1980; Edgington & Onghena, 2007; Gorman & Allison, 1997; Kazdin, 2011; Ludbrook, 1994; Onghena & Edgington, 2005; Dugard, File & Todman, 2012).

An alternative to these nonparametric tests is time series analysis, which investigates whether there is a significant change in level and/or slope between the phases. This technique is also suitable for the analysis of data when serial dependency is present. It, however, requires many data points to determine the existence and the pattern of autocorrelation and to identify the model correctly, which could cause problems for small-*n* experiments where the phases are usually rather short. Another difficulty is the complexity of the mathematical theories on which it is based (Box, Jenkins & Reinsel, 1994; Gorman & Allison, 1997; Kazdin, 2011). Another alternative is hierarchical linear modeling (HLM). Van den Noortgate and Onghena (2003a, 2003b) suggested using such a model for calculating effect sizes and for combining effect sizes of single-case data or combining the raw data of several studies. By modeling the hierarchical structure, the possible dependence of the scores is taken into account. Also study or case characteristics can be included as covariates to explain possible variation. For time series analysis a plug-in for the R commander already exists

(RcmdrPlugin.epack); this is not the case for hierarchical linear modeling, but R packages do exist (e.g., 'nlme').

As a useful extension of the RcmdrPlugin.SCDA package, these techniques could be included. Not only the HLM suggestion of Van den Noortgate and Onghena (2003a, 2003b), but also other methods for combining effect sizes could be adopted in the package. The simplest method is probably just taking the (weighted) average of all effect sizes to obtain one overall measure that reflects the general effect of the intervention over the different studies. This could for example be done with the effect size estimates already included in the package. Additionally, more effect size measures could be added, such as the percentage of all non-overlapping data (PAND; Parker, Hagan-Burke, & Vannest, 2007). Other possibilities for combining effect sizes include the three approaches presented by Busk and Serlin (1992), which differ in the assumptions made about the distribution and the variability of the data. Also the inclusion of more methods for combining p-values could be interesting: Stouffer's method, in which the Zs associated with the p-values are added and divided by the square root of the number of studies, after which the resulting Z-value is converted to an overall p; Mosteller and Bush's modification, that computes a t-test on the obtained Z value; Winer's suggestion of adding t-values and dividing the sum by the square root of the degrees of freedom (Rosenthal, 1978); and the iterative procedure for combining p-values, by applying more than one combining function to the same partial tests and then combining the resulting second order p-values into a third order of combination by means of a combining function, until the final p-value becomes reasonably invariant (Pesarin, 2001; Pesarin & Salmaso, 2010).

References

Barlow, D. H., Hayes, S. C., & Nelson, R. O. (1984). *The scientist practitioner: Research and accountability in clinical and educational settings.* New York: Pergamon.

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

Bliss, S. L., Skinner, C. H., Hautau, B., & Carroll, E. E. (2008). Articles published in four school psychology journals from 2000 to 2005: An analysis of experimental/intervention research. *Psychology in the Schools, 45,* 483-498.

Box, G. E. P., Jenkins, G. M., & Reinsel, G. C. (1994). *Time series analysis: Forecasting and control* (3rd ed.). Englewood Cliffs, NJ: Prentice Hall.

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

Bulté, I., & Onghena, P. (2009). Randomization tests for multiple baseline designs: An extension of the SCRT-R package. *Behavior Research Methods, 41,* 477-485.

Bulté, I., & Onghena, P. (2012). When the truth hits you between the eyes: A software tool for the visual analysis of single-case experimental data. *Methodology, 8,* 104-114.

Bulté, I., Van den Noortgate, W., Heyvaert, M., & Onghena, P. (submitted). Meta-analysis of single-case studies: The SCMA package. Manuscript submitted for publication.

Busk, P. L., & Serlin, R. C. (1992). Meta-analysis for single-case research. In T.R. Kratochwill & J.R. Levin (Eds.), *Single-case research designs and analysis: New directions for psychology and education* (pp. 187-212). Hillsdale, NJ: Lawrence Erlbaum.

Dugard, P., File, P., & Todman, J. (2012). *Single-case and small-n experimental designs: A practical guide to randomization tests* (2nd ed.). New York: Routledge Taylor & Francis Group.

Edgington, E. S. (1973). The random-sampling assumption in "Comment on component-randomization tests." *Psychological Bulletin, 80*, 84-85.

Edgington, E. S. (1980). Validity of randomization tests for one-subject experiments. *Journal of Educational Statistics*, **5**, 235-251.

Edgington, E. S., & Onghena, P. (2007). *Randomization tests* (4th ed.). London: Chapman & Hall/CRC.

Fox, J. (2005). The R Commander: A basic-statistics graphical user interface to R. *Journal of Statistical Software, 14.*

Fox, J. (2007). Extending the R commander by "plug-in" packages. *R News*, 7(3), 46-52. URL http://CRAN.R-project.org/doc/Rnews/.

Franklin, R. D., Allison, D. B., & Gorman, B. S. (Eds.). (1997). *Design and analysis of single-case research.* Mahwah, NJ: Lawrence Erlbaum Associates.

Gorman, B. S. & Allison, D. B. (1997). Statistical alternatives for single-case designs. In R. D. Franklin, D. B. Allison, & B. S. Gorman (Eds.), *Design and analysis of single-case research* (pp. 119-158). Mahwah, NJ: Erlbaum.

Hammond, D., & Gast, D. L. (2010). Descriptive analysis of single-subject research designs: 1983-2007. *Education and Training in Autism and Developmental Disabilities, 45,* 187-202.

Hornik, K. (2012). *The R FAQ: Frequently asked questions on R.* Retrieved February 17, 2013 from cran.r-project.org/doc/FAQ/R-FAQ.html.

Huber, P., & Ronchetti, E. (2009). *Robust statistics* (2nd ed.). Hoboken, NJ: John Wiley and Sons.

Kazdin, A. E. (2011). *Single-case research designs: Methods for clinical and applied settings* (2nd ed.). New York: Oxford University Press, 2010.

Kelley, K. (2007). Methods for the behavioral, educational, and social sciences: An R package. *Behavior Research Methods, 39*, 979-984.

Ludbrook, J. (1994). Advantages of permutation (randomization) tests in clinical and experimental pharmacology and physiology. *Clinical & Experimental Pharmacology & Physiology, 21,* 673-686.

Matson, J. L., Turygin, N. C., Beighley, J., & Matson, M. L. (2012). Status of single-case research designs for evidence-based practice. *Research in Autism Spectrum Disorders, 6,* 931-938.

Onghena, P., & Edgington, E. S. (2005). Customization of pain treatments: Single-case design and analysis. *Clinical Journal of Pain, 21,* 56-68.

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.

Pesarin, F. (2001). *Multivariate permutation tests. With applications in biostatistics.* New York: John Wiley & Sons.

Pesarin, F., & Salmaso, L. (2010). *Permutation tests for complex data: Theory, applications and software.* Chichester, UK: Wiley.

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.

Swaminathan, H., & Rogers, H. J. (2007). Statistical reform in school psychology research: A synthesis. *Psychology in the Schools, 44,* 543-549.

ter Kuile, M. M., Bulté, I., Weijenborg, P. T. M., Beekman, A., Melles, R., & Onghena, P. (2009). Therapist-aided exposure for women with lifelong vaginismus: A replicated single-case design. *Journal of Consulting and Clinical Psychology, 77,* 149-159.

Tukey, J. W. (1977). *Exploratory data analysis.* Reading, MA: Addison-Wesley.

Van den Noortgate, W., & Onghena, P. (2003a). Combining single-case experimental data using hierarchical linear models. *School Psychology Quarterly*, *18*, 325-346.

Van den Noortgate, W., & Onghena, P. (2003b). Hierarchical linear models for the quantitative integration of effect sizes in single-case research. *Behavior Research Methods, Instruments, & Computers*, *35*, 1-10.

Wilcox, R. R. (2005). *Introduction to robust estimation and hypothesis testing* (2nd ed.). San Diego, CA: Elsevier Academic Press.

Appendix: Getting started with the RcmdrPlugin.SCDA

Installation

To make the SCDA GUI work, R should be downloaded and installed. This can be done at no cost from the CRAN (Comprehensive R Archive Network; cran.r-project.org) website. Hornik (2012) gives a detailed explanation of how to do this for Windows, Macintosh and UNIX (the R package presented here is, however, only tested on Windows). Once R is running, the GUI can be installed from within the R console (Packages -> Install packages). After choosing a CRAN mirror nearby, the 'RcmdrPlugin.SCDA' package should be selected from the list. Note that the downloading process can take a while because several supporting packages are downloaded automatically as well. Then the RcmdrPlugin.SCDA package can be loaded into R by selecting 'Packages -> Load package' (or by typing the command *library('RcmdrPlugin.SCDA')* into the R console). Only this last step needs to be repeated when using the GUI. Note that under Windows, the R commander functions best with the single-document interface (SDI: Edit -> GUI preferences -> Single or multiple windows: SDI). By loading the RcmdrPlugin.SCDA package, the R commander opens with 'SCDA' as an additional menu button (Figure A1).

Figure A1. By loading the RcmdrPlugin.SCDA package, the "SCDA" menu button is added to the R commander interface.

Data Input

Most of the functions in the SCDA GUI need data input, which can be taken from a .txt file that has been created in advance in a text editor (e.g., EditPad or NotePad) or in Excel (save the file as 'Text(Tab delimited)'). Text files containing the raw data should consist of two tab-separated columns for single-

case phase and alternation designs: the first with the condition labels ("A" and "B" when there are two conditions/phases, and "A1", "B1", "A2" and "B2" for three or four phases) and the second with the observations. For multiple-baseline designs they should consist of these two columns for each unit. This way, each row represents one measurement occasion. Text files containing p-values for the 'combine' function should consist of one column of p-values. In text files containing the possible start points for multiple baseline designs, each row should contain all possibilities for one unit, separated by a tab. It is important not to label the rows or columns.

Another way of data input is by using the active data set of the R commander. Creating an active data set has the advantage that also other functions than the SCDA-functions (e.g., built-in statistical functions of the R commander or from other plug-ins) can be applied to the data. This way one can fully benefit from the functionality of the R commander and its plug-ins. There are several ways to construct this active data set: a text file created as described above can be read into the R commander via the Data menu (Figure A2: Data -> Import data -> from text file, clipboard, or URL -> uncheck the box 'Variable names in file' if necessary -> OK -> select the text file within the 'Open file' dialog box), or data can be entered directly via 'Data -> New data set'. The active data set can be consulted at any time by clicking the 'View data set' button.
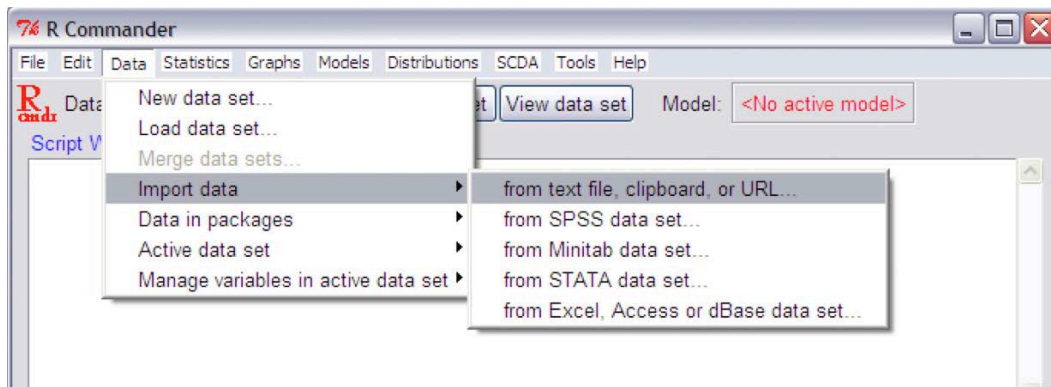

Figure A2. Constructing the active data set by reading a text file into the R commander.

Functions

The SCDA menu contains three submenus with several menu items. Each menu item opens a dialogue box.

SCVA (Single-Case Visual Analysis)

Graphical display   The observed single-case data are plotted. As can be seen in Figure A3, two selections should be made in the dialogue box. In section A, the *design type* used has been indicated. The options are:

Phase Designs
Comparisons are made within a time series and the subject's performance is evaluated over time across baseline (A) and intervention (B) phases.

AB Phase Design	*ABA Phase Design*:	*ABAB Phase Design*
All baseline measurements precede all treatment measurements.	Withdrawal or reversal design in which the treatment is administered between two baseline phases.	An extra intervention phase is added.

Alternation Design
The basic strategy is the rapid alternation of two or more conditions within a single subject.

Completely Randomized Design	*Alternating Treatments Design*	*Randomized Block Design*
The random assignment procedure mirrors the one used when randomly assigning subjects to different groups for an independent samples *t* test (no restrictions).	The temporal clustering of treatments is prevented by ensuring that the randomization does not permit more than a specified number of successive time blocks with the same condition.	Adjacent treatment times are grouped together in blocks and the conditions are assigned randomly within each block.

467

Multiple Baseline Design
This is an extension of the basic AB phase designs, in which several of those AB designs are implemented simultaneously to different persons, behaviors, or settings. A characteristic feature is that the intervention is introduced in a staggered way to the different units.

In section B, the *data file* in which the data can be found should be selected. This could either be the active data set or a previously created text file. In the latter case, the file can be chosen with the 'Select File' button.
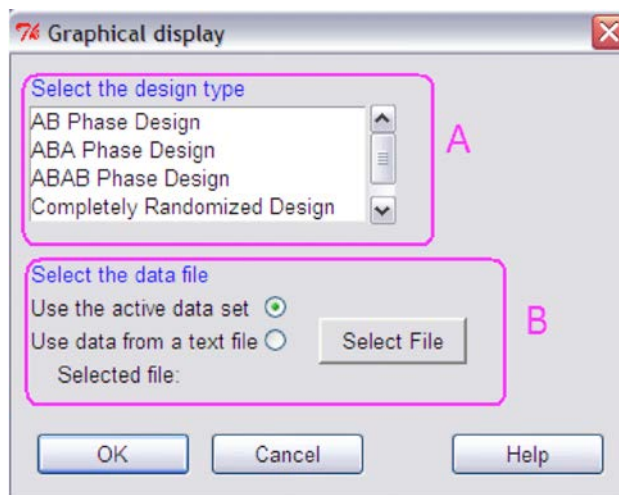
Figure A3. The dialogue box for the 'graphical display' menu.

Plot measure of central tendency   a measure of central tendency is plotted as a horizontal reference line superimposed on the raw time series data.
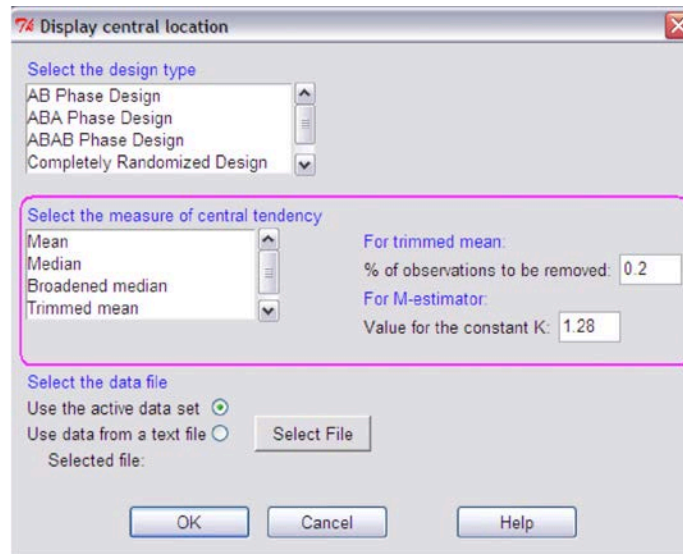
Figure A4. The dialogue box for the 'Plot measure of central tendency' menu.

In the marked section of the dialogue box (see Figure A4), the *measure of central tendency* should be selected. There are four built-in possibilities:

Mean
The arithmetic mean (the average) of the data. This measure is rather sensitive to outliers.

Median
The middle value of the data (for an even number of data points, the median equals the average of the two central data points). This measure is more robust than the mean, but it only takes into account the central data points, while disregarding other numerical information.

Broadened median
Calculated based on the three, four, five or six middle values of the data set (depending on the total number of data points). This measure is more robust than the mean, and sensitive to a larger proportion of the data than the median.

Trimmed mean
Calculates the mean after discarding the extreme observations. The percentage of observations that has to be removed from each end of the distribution can be indicated in the upper right box (Figure A4). This can be any value from 0 (= regular arithmetic mean) to 0.5. The default value is 20 percent. The trimmed mean is more robust than the mean and less affected by observations in the centre of the distribution than the (broadened) median.

M-estimator
Huber's M-estimator of location first evaluates each observation to determine if it is actually an outlier compared to the rest of the data and then gives less weight to those outlying values. For this evaluation a constant K needs to be specified that can have any value between 0 and ∞. Usually a percentile of the standard normal distribution is chosen. Wilcox (2005) suggests using $K = 1.28$, which corresponds to the 90th percentile of the standard normal distribution and covers 80 percent of the underlying distribution. By determining the sensitivity of the estimator, one can balance between robustness and efficiency (see e.g., Huber & Ronchetti, 2009).

Plot estimate of variability   Information about variability in the data is displayed by one of three methods.
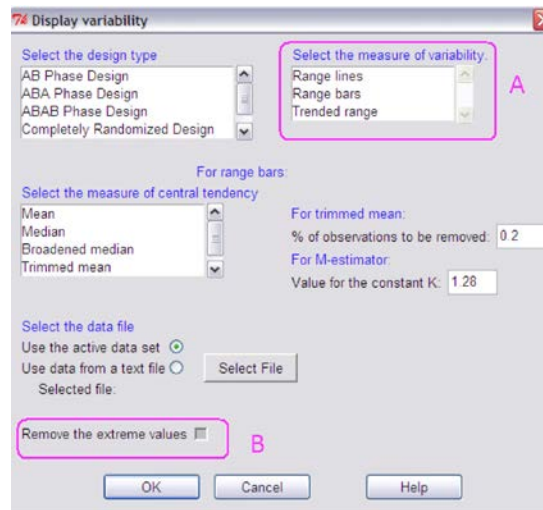

Figure A5. The dialogue box for the 'Plot estimate of variability' menu.

In box A (Figure A5) the *measure of variability* should be selected. The three options are:

Range lines
Range lines are a pair of lines parallel to the X-axis, passing through the lowest and highest values for each phase, and superimposed on the raw data.

Range bar graphs
Range bar graphs consist of a vertical line for each phase, connecting the minimum and the maximum value, with a small horizontal bar crossing this line to display a measure of central location ((trimmed) mean, (broadened) median, or M-estimator). This estimate of central tendency should be selected in the middle part of the dialogue box (Figure A5).

Trended ranges
Trended ranges display changes in variability within phases by two lines, one connecting the minimum values of the phase halves and one connecting the maxima.

For all these methods the *influence of outliers* may be lessened by using a trimmed range, in which only a sample of the data set is used. This can be selected in box B (Figure A5): default the whole dataset is used, but by checking the box the 10 to 20 percent extreme values are removed from each phase.

Plot estimate of trend   This function visualizes systematic shifts in central location over time using several methods. The method of *trend visualization* should be selected (Figure A6):

Vertical line plot
A *vertical line plot* draws the deviations from each data point to a measure of central tendency against time. The measure of central tendency should be selected in the middle part of the dialogue box (Figure A6).

Trend lines

Trend lines superimpose a linear function on the raw data, which shows if there is an increase or a decrease in the observed behavior over time.

Least squares regression	*Split-middle method*	*Resistant trend line fitting*
Minimizes the squared vertical distances between the regression line and the data points.	Connects the crossings of the median dependent variable value and the middle time value of both phase halves	Comparable to the split-middle method, but here the phases are divided into three sections instead of two. This makes this method more suited for longer time series.

Running medians

The presence of a nonlinear trend can be displayed with *running medians*, with which the time series is smoothed by dividing it into successive segments of a given size and calculating the median for each segment (Tukey, 1977). Three sizes of segments are easy to use with time series data: *batch size 3, batch size 5, and batch size 4 averaged by pairs*.
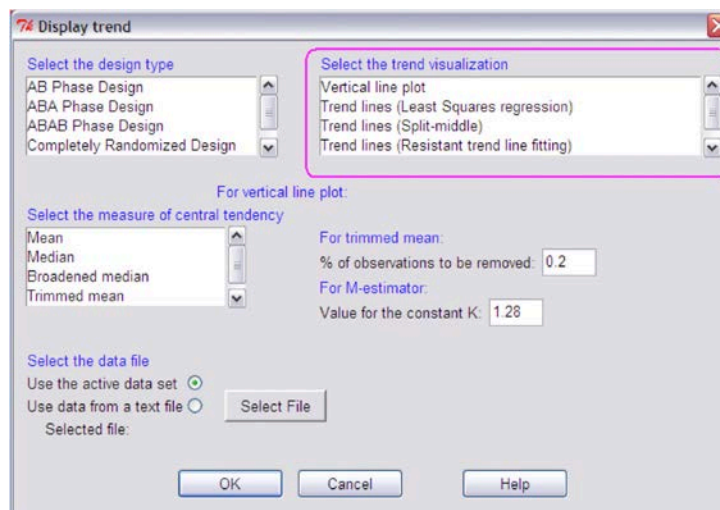

Figure A6. The dialogue box for the 'Plot estimate of trend' menu.

SCRT (Single-Case Randomization Tests)

Design your experiment

Number of possible assignments

The number of assignment possibilities for the specified design is calculated. In box A (Figure A7) the *total number of observations* in the experiments should be indicated. This is not necessary when the design used is a multiple baseline design. In box B a *constraint* on the randomization schedule should be provided. In phase designs the moment of phase change is randomly determined, so the restriction is placed on the minimum number of observations per phase. In alternation designs the observations are randomly assigned to different conditions, so in other words the treatment order is randomly determined. To avoid too long sequences of the same condition, in alternating treatment designs there is a restriction on the maximum number of consecutive administrations of the same condition. For multiple baseline designs, there is one extra selection to be made: in box C the location of the file with the *possible start points* should be selected (see 'Data input').

Figure A7. The dialogue box for the 'Number of possible assignments' menu.

Display all possible assignments

All assignment possibilities for the specified design are enumerated. In the dialogue box (Figure A8) can be chosen if the possibilities should only be displayed in the output window of the R commander, or also should be saved to a file. The location to save the assignment possibilities has to be indicated with the 'Select location' button.
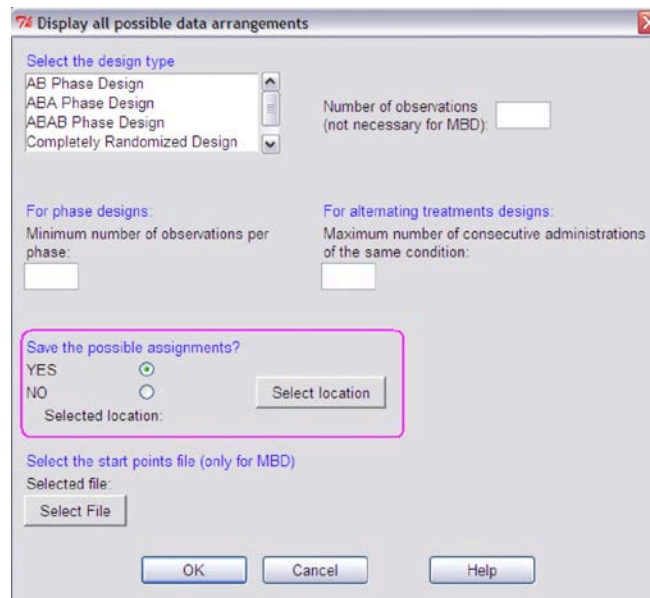

Figure A8. The dialogue box for the 'Display all possible assignments' menu.

Choose 1 possible assignment

One assignment possibility is randomly selected from all theoretical possibilities. The dialogue box is similar to that of the 'Number of possible assignments' function.

Analyze your data

Observed test statistic

The observed test statistic is calculated from the obtained raw data. There are several built-in possibilities as *test statistic* (see Figure A9). For alternation designs, multiple-baseline designs and AB phase designs, there are three

474

options: *"A-B", "B-A",* and *"/A-B/"*, which stand for the (absolute value of the) difference between the condition/phase means. For phase designs with more than two phases, six more options are available: *"PA-PB", "PB-PA",* and *"/PA-PB/* refer to the (absolute value of the) difference between the means of phase means, and *"AA-BB", "BB-AA",* and *"/AA-BB/"* represent the (absolute value of the) difference between the sums of phase means.
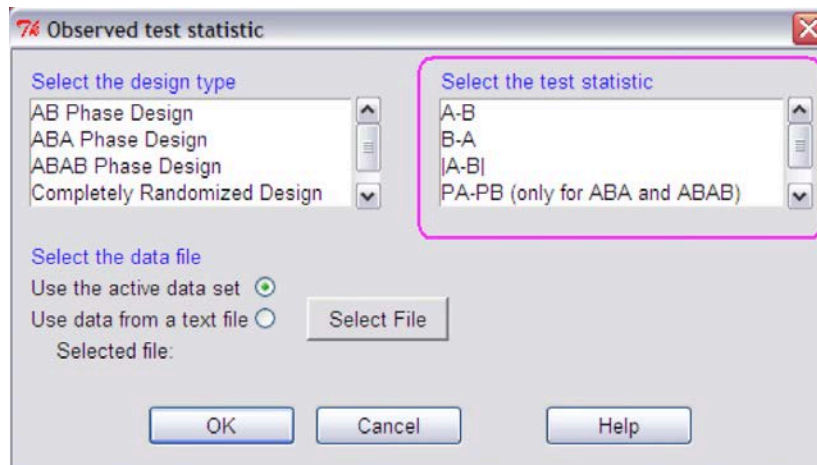

Figure A9. The dialogue box for the 'Observed test statistic' menu.

Randomization distribution

The randomization distribution is generated. One can choose between the exhaustive ('systematic') and the nonexhaustive ('Monte Carlo') randomization distribution (Figure A10 box A). For the exhaustive randomization distribution all assignment possibilities are enumerated, while the nonexhaustive randomization distribution is generated by a random sample of all assignment possibilities. The size of this random sample should be indicated in the box 'number of randomizations'. In box B (Figure A10) can be chosen if the randomization distribution should only be displayed in the output window of the R commander, or also should be saved to a file. The location to save the randomization distribution has to be indicated with the 'Select location' button
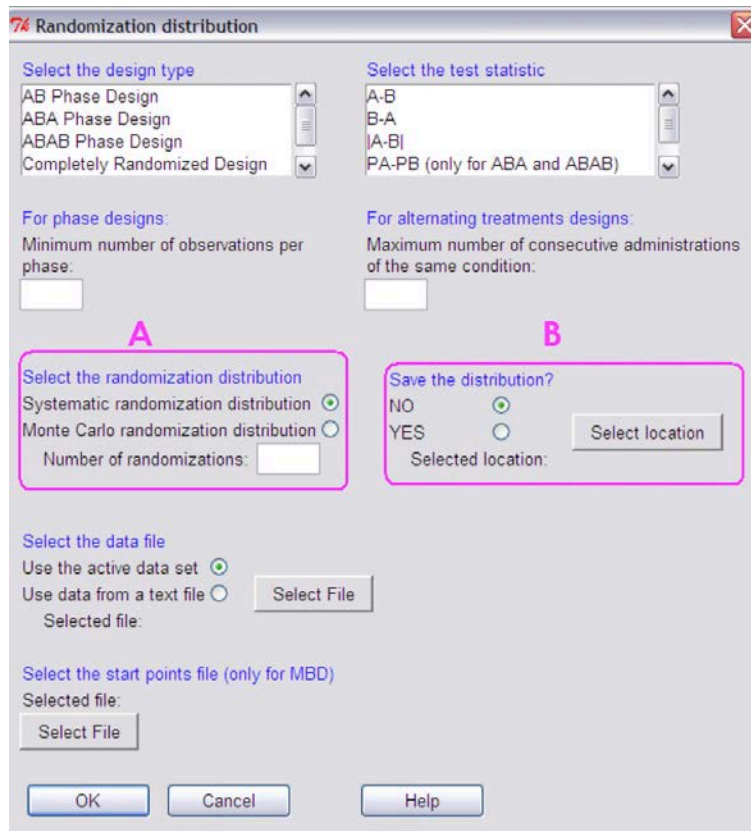
Figure A10. The dialogue box for the 'Randomized distribution' menu.

P-value
The *p*-value corresponding to the observed value of the test statistic is obtained by locating this value in the exhaustive or nonexhaustive randomization distribution. The dialogue box is similar to that of the 'randomization distribution' function.

SCMA (Single-Case Meta-Analysis)

Calculate effect size The specified effect size measure is calculated (Figure A11). Four effect size measures are included in the RcmdrPlugin.SCDA:

Standardized mean difference
The mean of the baseline condition is subtracted from the mean of the treatment condition, and this difference is divided by the baseline standard deviation.

Pooled standardized mean difference
Similar to the standardized mean difference, but with the pooled standard deviation used as denominator

PND (percentage of nonoverlapping data)
The percentage of data points in the treatment phase that exceed the most extreme value in the baseline phase (i.e., the proportion lower than the lowest baseline point for interventions designed to decrease unwanted behavior, and the proportion higher than the highest baseline point for interventions intended to increase desired behavior).

PEM (percentage of data points exceeding the median)
The overlap is calculated as the percentage of data points in the treatment phase that are located above the extended median line of the previous baseline phase (or below this line, if the undesired behavior is expected to decrease).
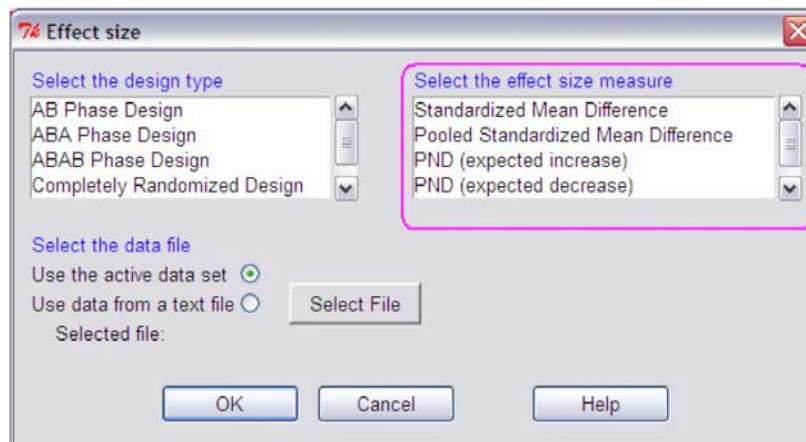

Figure A11. The dialogue box for the 'Calculate effect size' menu.

Combine p-values   Nonparametric combination of p-values, with the multiplicative approach using Pearson's formula or the additive approach using Edgington's formula (Figure A12 box A). In box B (Figure A12) the location of the file in which the p-values to be combined can be found should be selected.
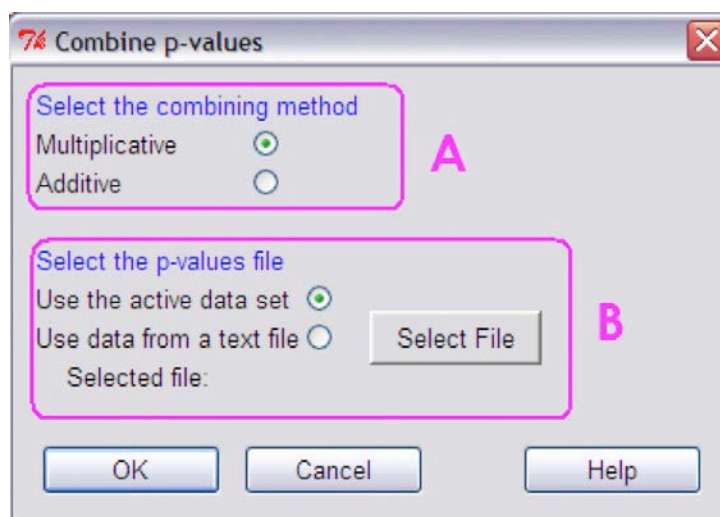

Figure A12. The dialogue box for the 'Combine *p*-values' menu.