

# Research on Social Work Practice

<http://rsw.sagepub.com>

---

## **Pediatric Pain and Anxiety: A Meta-Analysis of Outcomes for a Behavioral Telehealth Intervention**

Gary Holden, David J. Bearison, Diane C. Rode, Merri Fishman-Kapiloff, Gary Rosenberg and Patrick Onghena

*Research on Social Work Practice* 2003; 13; 693

DOI: 10.1177/1049731503254107

The online version of this article can be found at:  
<http://rsw.sagepub.com/cgi/content/abstract/13/6/693>

---

Published by:

 SAGE Publications

<http://www.sagepublications.com>

**Additional services and information for *Research on Social Work Practice* can be found at:**

**Email Alerts:** <http://rsw.sagepub.com/cgi/alerts>

**Subscriptions:** <http://rsw.sagepub.com/subscriptions>

**Reprints:** <http://www.sagepub.com/journalsReprints.nav>

**Permissions:** <http://www.sagepub.com/journalsPermissions.nav>

**Citations** (this article cites 15 articles hosted on the SAGE Journals Online and HighWire Press platforms):  
<http://rsw.sagepub.com/cgi/content/refs/13/6/693>

# **Pediatric Pain and Anxiety: A Meta-Analysis of Outcomes for a Behavioral Telehealth Intervention**

**Gary Holden**

*School of Social Work, New York University*

**David J. Bearison**

*The Graduate School & University Center,  
City University of New York*

**Diane C. Rode**

**Merri Fishman-Kapiloff**

**Gary Rosenberg**

*Mount Sinai School of Medicine, Mount Sinai Hospital*

**Patrick Onghena**

*Katholieke Universiteit Leuven*

*Objectives: This study sought to explore the utility of using single system designs (SSD) in an optimal set of circumstances. The effects of a computer network intervention on three outcomes (pain intensity, pain aversiveness, and anxiety) were examined. Method: Meta-analysis of three studies using restricted alternating treatment designs was conducted. Results: Children reported significantly less pain intensity, pain aversiveness, and anxiety in the SBW condition than in the GPM condition. Both sex and age moderated these effects. Conclusions: An SSD/meta-analysis approach can provide multiple views of the impact of an intervention.*

**Keywords:** *single subject design; meta-analysis; pain; anxiety; technology*

The discussion of the single system design (SSD) for assessing the effectiveness of interventions in social work reached a high point in the June 1996

---

**Authors' Note:** The authors would like to acknowledge Jaime Rosensweig, John Ward, Gilit Ullmann, Karen Marcinczyk, and Dr. Fredrick Suchy, who made this project possible. Also, they would like to thank Alexandru Cuc for his statistical support. Finally, they would like to thank the STARBRIGHT Foundation, the Mount Sinai Children's Center Foundation, and the Mount Sinai Medical Center for their support in carrying out this research. Correspondence may be addressed to Gary Holden, New York University, ESSW, Room 407, 1 Washington Square North, New York, NY 10003; e-mail: gary.holden@nyu.edu.

Research on Social Work Practice, Vol. 13 No. 6, November 2003 693-704  
DOI: 10.1177/1049731503254107  
© 2003 Sage Publications

issue of *Social Work Research*. Consider two contrasting points of view presented there. According to Thyer (1996),

- The incorporation of SSDs into social work practice to date has been a significant advance.
- Training social workers to use these designs is now a common component of social work education.
- A substantial literature exists that demonstrates the use of these designs in practice.
- SSDs are flexible and relatively easy-to-use evaluation tools.

Wakefield and Kirk (1996) countered,

- A-B designs do not allow strong causal inferences regarding interventions.
- The use of rapid assessment instruments in SSDs may not be a cost-effective approach.
- The idea that aggregation of results of multiple SSDs done by practitioners will result in findings that have external validity is not tenable.

We will not address meta-theory or ideology related to this debate nor will we review the standard approaches to SSD research that have been covered so well elsewhere (e.g., Bloom, Fischer, & Orme, 1999; Witkin, 1996). Rather, we report on a unique set of design decisions made in response to a specific evaluation problem (the need to assess the impact of a computer network intervention). This problem, combined with the emergence of new methodological choices provided us with an opportunity to try out some ideas that had previously been discussed primarily in the abstract (e.g., Corcoran, 1985, 1986; Gingerich, 1984; Videka-Sherman, 1986).

Jayaratne, Tripodi, and Talsma (1988) moved beyond the discussion of aggregating SSD in their study. Their sample included 42 cases from a family service agency for which they had SSDs for 50 problems. The SSDs were analyzed using three different procedures: Shewart charts, a binomial method, and visual analysis of the data by three independent raters. The results of each SSD were categorized into one of three change groups (significant positive change, significant negative change, no significant change/no change).

Jayaratne et al. (1988) aggregated the SSD data using two methods. First, they computed the proportion of SSDs falling into each of the change groups. Next, they computed effect sizes according to Gingerich's (1984) recommendations. They concluded in part that

regardless of the method of analysis, the social workers in the agency studied apparently fared less well in addressing marital problems than in dealing with

child management and individual problems. This consistency in findings provides evidence to both the social workers and administrators that a specific problem exists. (Jayaratne et al., 1988, p. 125)

One of the advantages of using meta-analysis to cumulate the results across replicated SSDs is the increased power in testing the null hypothesis that the treatment is ineffective. It is even possible that none of the individual  $p$  values is smaller than .05 but that the combined  $p$  value is. This means that there is no statistical evidence that the treatment is effective when looking at each of the experiments separately but that the combination of experiments shows there are effective ingredients and therefore that the intervention should not be abandoned without further notice (Onghena, 1994).

This kind of situation is similar to the so-called misguided "failures to replicate" described by Rosnow and Rosenthal (1989). Consider a researcher finding an experimental effect with associated  $p$  value of .02 and therefore rejecting the null hypothesis. Another researcher tries to replicate this finding and cannot reject the null hypothesis because his obtained  $p$  value is .11. Is this a failure to replicate? It is only when following a blind accept/reject dichotomous testing rule for each study separately. It could well be that the effect sizes of both studies are identical and that the difference between the  $p$  values of the studies is directly related to their numbers of observations. A  $p$  value of .11 could therefore represent a perfect corroboration.

To illustrate the perspective of meta-analysis, consider an additional three independent researchers finding  $p$  values of .47, .14, and .39, respectively. A  $p$  value less than .50 means that the effect is in the right direction. What is the probability that five independent tests result in an effect in the right direction? Following the binomial formula, this is .03125. So the combined  $p$  value of .02, .11, .14, .39, and .40 is .03 even with this simplified sign test.

The goal of the present study was to determine if given a group of SSDs, could the results be combined in ways that would provide insight regarding the overall group of SSDs and subsets of SSDs within that group? This approach may help social workers to consider how much use SSDs would have for practice if they

- were examining interventions that have short acting effects, thereby eliminating the methodological and ethical concerns regarding the use of withdrawal designs;
- could move beyond visual analysis of results and easily employ null hypothesis statistical testing to assist in decision making regarding the importance of results; and
- could employ statistical software to aggregate the results across entire series of SSDs and within subsets of the entire series.

The goal of this article is to explore the utility of using meta-analysis with SSDs. We attempt to accomplish this goal through the presentation of a series of three evaluations of STARBRIGHT World (SBW) that have been reported individually elsewhere (Holden et al., in press; Holden, Bearison, Rode, Rosenberg, & Fishman, 1999, 2000). Details of those results are repeated here. In addition, new meta-analyses are presented that combine results from all three evaluations and explore potential moderating variables.

## METHOD

SBW is a computer network that allows children hospitalized throughout the United States to communicate with each other. SBW consists of a variety of interactive components designed to promote communication, distraction, information, self-expression, and social support. Children can communicate with each other via video-conferencing, chat rooms, bulletin boards, and e-mail. They can also access health care-related information and play games. The availability of SBW has increased from five hospitals in 1995 to 88 hospitals in May of 2001 (Additional background regarding SBW is available at [www.starbright.org](http://www.starbright.org)).

### Participants

In the three evaluations combined, there were a total of 92 participants between the ages of 7 and 19 hospitalized on pediatric services in an urban teaching hospital. The average age was 12.09, and 54.3% of the sample was female.

### Measures

Self-report, visual analog scales were used to obtain ratings of the children's perceptions of pain intensity, pain aversiveness, and anxiety (McGrath, 1990; McGrath et al., 1996; Tuckman, 1988). Self-report measures provide reliable and valid assessments of pain and can be used with children older than four (Agency for Health Care Policy and Research, 1992).

In general, visual analog scales have demonstrated good psychometric properties and are potentially sensitive to weak intervention effects (Eggebrecht, Bautz, Brenig, Pflingsten, & Franz, 1989; Eimer & Freeman, 1998; Varni, Blount, Waldron, & Smith, 1995).

The timing of observations was conceptualized as an ecological momentary assessment approach (e.g., Affleck et al., 1998; Litt, Cooney, & Morse,

1998; Shiffman & Stone, 1998), in which children were asked to rate their current level of pain intensity, pain aversiveness, and anxiety.

### **Interventions**

There were two conditions—an experimental and a contrast condition. In the experimental condition (SBW), the child had the opportunity to use the various components of the computer network (e.g., video-conferencing, chat rooms, bulletin boards, e-mail, games, health education modules). The contrast condition was referred to as the General Pediatric Milieu (GPM). In the GPM, the child engaged in pursuits such as playroom group activity, an art activity, individual therapeutic play activity, family interaction, visits from family and friends, eating, watching TV or video, reading, or talking on the telephone. These are typical kinds of activities for children in pediatric units.

### **Procedure**

A restricted alternating treatment design (RATD) (e.g., BAAABBABBBBAB and so on) was used. This experimental SSD allowed us to determine whether the SBW condition (B phase) had significantly greater effects for the hospitalized child than the GPM condition (A phase). In an RATD, treatments are randomly assigned to time periods and an upper limit to the number of consecutive periods for a treatment is imposed (Onghena & Edgington, 1994). In each of the three evaluations, this upper limit was six.

### **Hypotheses**

Each of the three evaluations hypothesized that children would experience significantly less pain intensity, pain aversiveness, and anxiety when in the SBW condition. This meta-analysis tested these hypotheses using data from all three evaluations and explored the influence of age and sex as moderator variables.

## **RESULTS**

Differences between the conditions were tested for each individual ( $n = 92$ ) for each of the three outcomes using randomization tests of the mean difference between the conditions (Onghena & Edgington, 1994). It was found that 12.3% (34/276) of the individual level contrasts were statistically significant at an alpha level of .05. Yet only 2.9% (8/276) of the contrasts were

**TABLE 1: Meta-Analytic Effects for Females Versus Males**

	n <sup>a</sup>	Age	Pain Intensity	Pain Aversiveness	Anxiety
Females					
Mean	16.9	12.4	.38	.04	.31
n <sup>b</sup>			41	43	40
p for meta-analysis <sup>c</sup>			p = .0002	p = .0023	p = .0002
Males					
Mean	18.1	11.8	.19	.05	.32
n <sup>b</sup>			32	39	36
p for meta-analysis <sup>c</sup>			p = .1456	p = .0001	p = .0002

a. Mean number of observation points.

b. Number of participants (reduced *ns* due to participants who reported no variation across observations).

c. Alpha level = .00833.

**TABLE 2: Meta-Analytic Effects for Age Groups**

	n <sup>a</sup>	Age	Pain Intensity	Pain Aversiveness	Anxiety
Age 7 to 10					
Mean	17.3	8.9	.17	.05	-.01
n <sup>b</sup>			21	27	23
p for meta-analysis <sup>c</sup>			p = .1641	p = .0157	p = .2107
Age 11 to 13					
Mean	17.1	11.7	.47	.06	.48
n <sup>b</sup>			27	32	30
p for meta-analysis <sup>c</sup>			p = .0003	p < .0001	p = .0007
Age 14 to 19					
Mean	18.1	16.1	.22	.02	.42
n <sup>b</sup>			25	23	23
p for meta-analysis <sup>c</sup>			p = .0868	p = .1015	p = .0059

a. Mean number of observation points.

b. Number of participants (reduced *ns* due to participants who reported no variation across observations).

c. Alpha level = .00555.

significant at the more appropriate Bonferroni analysiswise alpha level of .00018.

Bonferroni analysiswise alpha levels were used in these and the analyses presented in Tables 1 to 3. That means that the overall alpha level is equal to .05 for the analyses presented in a given table. The reader should bear in mind that power decreases in these analyses because of the reduced *n* of

**TABLE 3: Meta-Analytic Effects for Sex by Age Groups**

	<i>n</i> <sup>a</sup>	Age	Pain Intensity	Pain Aversiveness	Anxiety
Females, 7-10					
Mean	15.0	8.9	.30	.05	.17
<i>n</i> <sup>b</sup>			11	12	11
<i>p</i> for meta-analysis <sup>c</sup>			<i>p</i> = .0835	<i>p</i> = .4725	<i>p</i> = .0851
Females, 11-13					
Mean	18.6	11.8	.56	.05	.29
<i>n</i> <sup>b</sup>			14	18	15
<i>p</i> for meta-analysis <sup>c</sup>			<i>p</i> = .0009	<i>p</i> = .0009	<i>p</i> = .0067
Females, 14-19					
Mean	16.7	15.6	.28	.02	.43
<i>n</i> <sup>b</sup>			16	13	14
<i>p</i> for meta-analysis <sup>c</sup>			<i>p</i> = .0466	<i>p</i> = .0755	<i>p</i> = .0132
Males, 7-10					
Mean	19.3	8.9	.04	.05	-.19
<i>n</i> <sup>b</sup>			10	15	12
<i>p</i> for meta-analysis <sup>c</sup>			<i>p</i> = .5129	<i>p</i> = .0021	<i>p</i> = .5787
Males, 11-13					
Mean	15.4	11.6	.38	.09	.67
<i>n</i> <sup>b</sup>			13	14	15
<i>p</i> for meta-analysis <sup>c</sup>			<i>p</i> = .0484	<i>p</i> = .0018	<i>p</i> = .0217
Males, 14-19					
Mean	20.6	16.6	.10	.01	.40
<i>n</i> <sup>b</sup>			9	10	9
<i>p</i> for meta-analysis <sup>c</sup>			<i>p</i> = .4869	<i>p</i> = .3869	<i>p</i> = .1081

a. Mean number of observation points.

b. Number of participants (reduced *ns* due to participants who reported no variation across observations).

c. Alpha level = .00277.

participants. Use of a Bonferroni adjustment means that our conclusions based on these tests are quite conservative.

Meta-analyses were then used to combine the findings from the individual SSDs (e.g., Allison & Gorman, 1993; Salzberg, Strain, & Baer, 1987; Scruggs, Mastropieri, & Casto, 1987; White, 1987; White, Rusch, Kazdin, & Hartmann, 1989; Zucker et al., 1997). The Single Case Randomization Test software package was used for randomization tests and meta-analyses (Onghena, 1992, 1994; Onghena & Edgington, 1994; Van Damme & Onghena, 1995). SPSS for Windows 10.0 was used for additional analyses. The following meta-analytic findings have been reported elsewhere (Holden et al., 1999, 2000, 2002):



- *Evaluation 1*: Results favoring SBW for all three outcomes (mean difference between conditions [ $p$  obtained]): pain intensity (.29 [.042]), pain aversiveness (.04 [.025]), and anxiety (.34 [.09]).
- *Evaluation 2*: Results favoring SBW for all three outcomes: pain intensity (.34 [.0002]), pain aversiveness (.05 [.0018]), and anxiety (.39 [.0002]).
- *Evaluation 3*: Results favoring SBW for all three outcomes: pain intensity (.13 [.528]), pain aversiveness (.04 [.025]), and anxiety (.10 [.16]).

A Bonferroni analysiswise adjustment was used to set the alpha level at .05 for each of the series of three analyses in each evaluation. Although all findings favored the SBW condition, only three of nine contrasts were statistically significant ( $p < .01666$ ).

The next step was to aggregate the individual-level data from the three studies. These meta-analytic results were all statistically significant (alpha = .01666) and therefore support the a priori hypotheses. Children reported significantly less pain intensity, pain aversiveness, and anxiety in the SBW condition than in the GPM condition.

In addition to summarizing overall effects, meta-analysts typically seek to test the impact of moderating variables. Tables 1 to 3 examine the effects of respondents' sex and age on the three outcomes. In Table 1, the results favored the SBW condition and were statistically significant (alpha = .00833) in five of six instances. Yet, for pain intensity, the effect for males was one half the size of the effect for females and was not significant. Table 2 indicates that the statistically significant effects (alpha = .00555) were observed for all three outcomes but only for the 11- to 13-year-olds.

In Table 3, 4 of 18 effect sizes reached significance (alpha = .00277). Females in the 11- to 13-year age group reported significantly less pain intensity and pain aversiveness in the SBW condition than in the GPM condition. Males in both the 7- to 10-year and the 11- to 13-year age groups reported significantly less pain aversiveness in the SBW condition than in the GPM condition.

## DISCUSSION AND APPLICATIONS TO PRACTICE

The three studies combined in the present meta-analysis contrasted the effects of SBW with the GPM on children's perceptions of pain intensity, pain aversiveness, and anxiety. Although only 2.9% of the contrasts at the level of an individual child were statistically significant, the overall effects across studies were statistically significant for all three outcome variables. Moderator analyses revealed that the effects were slightly stronger for females than males and strongest for 11- to 13-year-olds. The strongest

effects when considering sex and age were for 11- to 13-year-old females on pain intensity and pain aversiveness and for 7- to 13-year-old males on pain aversiveness.

Some of the limitations of the present findings are that these studies involved convenience samples of children. They were conducted in a single medical center, by a single group of investigators, using a single set of outcome measures during three unique time periods. The external validity of the findings may be limited by any of these factors.

Let us consider three of the points raised earlier. How much utility would SSDs have for practice if the following occurred?

- We could use withdrawal designs more frequently.
- We could more easily employ null hypothesis statistical testing via a software program to assist in decision making regarding the importance of results.
- We could aggregate the results across entire series of SSDs and within subsets of the entire series.

In these evaluations, we were not only able to use withdrawal designs but were also able to randomly assign the treatments to times. What proportion of social work interventions would likely allow for such an approach? Given the type of problems most social work practitioners encounter, one might assume this proportion would be small. Yet one of us (Holden) is considering a study that would employ this design. This study would compare two supervision methods: observation of social work interns via videotaped sessions and supervision using process recordings. The outcomes would be supervisors' and supervisees' satisfaction and perceived utility of each supervisory session. Type of supervisory approach would be randomly assigned to weeks throughout the semester. Although there may be great utility in the use of randomized restricted alternating treatment designs in specific instances such as this, the small number of these instances means that the utility of this approach for social work in general is probably low.

In the present study, small effect sizes meant that visual analysis might not have been optimal for decisions about the effects. Although other statistical procedures are available and have been used with many SSDs, the present study provides another example of an additional technique that can be employed by social work researchers employing SSDs. What do we mean by small effects here? Although many of the findings reached statistical significance, the size of the effects was small. Although we know of no conventional effect size estimates that are applicable to this type of design (e.g., PVE), we can provide the reader with an alternative, using the pain intensity outcome as an example. The overall difference between the two conditions

was .297 on a 10-point scale. In other words, the difference between the two conditions was 3% of the range of the scale.

Discussion of the third issue flows directly from the discussion of the first two. Although the ability to create overall and subgroup effect sizes and test their significance was useful for our research team, this was a special set of circumstances. It remains to be seen whether future methodological developments will present the opportunities to apply these more rigorous designs to a broader spectrum of social work research problems. In addition, it remains to be seen whether such developments will be able to be applied by master's-level practitioners with limited research training, limited time, and little extrinsic motivation to conduct such research. Practitioners with extensive research training, some dedicated time, and both intrinsic and extrinsic motivation carried out the research described here.

This issue has always been the Achilles heel behind the recommendation that practitioners engage in the ongoing use of SSDs. Have large numbers of practitioners ever had the motivation, skills, institutional support, and external rewards to heed such a recommendation? The limited adoption of these SSDs by practitioners (Thyer, 1996) seems to indicate that the answer is no.

Perhaps a more modest recommendation would achieve greater success. Guyatt et al. (1990) recommended an n-of-1 consultation and tutorial service for clinical physicians that might be useful for social workers. That is, academic researchers skilled in SSDs could partner with practitioners to use SSDs in specific practice situations to assess interventions of interest. Rather than a single practitioner attempting to evaluate all or most cases, one or more practitioners could collaborate with an academic researcher in a small-scale, practice-based study.

Guyatt et al. (1990) reported on 73 n-of-1 trials that they carried out with community physicians. They conclude that

we believe that our results show that n-of-1 trials are feasible to conduct in clinical practice and often result in clinically important changes in clinicians' confidence and the management decisions [treatment decisions] themselves. . . . The methodology of the n-of-1 trial provides physicians with a set of tools that can further increase the scientific rigor of their clinical practice and increase the likelihood that the treatments they prescribe are indeed those that are best for the patient. (Guyatt et al., 1990, p. 298)

It is time for us to move beyond the notion of the practice researcher who is a skilled clinician and researcher. When one considers the level of knowledge and skills required in either pursuit, the plausibility of competence as a combined professional is called into question. Practice-based research is a complex endeavor requiring collaboration.

An SSD/meta-analysis research approach allows for multiple views of the impact of an intervention. One may obtain

- graphic views of intervention effects over time at the individual level;
- individual-level effect size estimates and tests of significance; and
- overall and subgroup effect size estimates across individuals (or studies) and tests of the significance.

This ability to aggregate across individuals is a partial solution to the lack of external validity of single subject designs. The applicability of this approach to other practice-based research problems remains an empirical question.

## REFERENCES

- Affleck, G., Tennen, H., Urrows, S., Higgins, P., Abeles, M., & Hall, C. (1998). Fibromyalgia and women's pursuit of personal goals: A daily process analysis. *Health Psychology, 17*, 40-7.
- Agency for Health Care Policy and Research. (1992). *Acute pain management: Operative or medical procedures and trauma* (AHCPR Publication No. 92-0032). Rockville, MD: USDHHS, Public Health Service, Agency for Health Care Policy and Research.
- Allison, D. B., & Gorman, B. S. (1993). Calculating effect sizes for meta-analysis: The case of the single case. *Behaviour Research and Therapy, 31*, 621-631.
- Bloom, M., Fischer, J., & Orme, J. G. (1999). *Evaluating practice: Guidelines for the accountable professional* (3rd ed.). Boston: Allyn & Bacon.
- Corcoran, K. J. (1985). Aggregating the idiographic data of single-subject research. *Social Work Research and Abstracts, 21*, 9-12.
- Corcoran, K. J. (1986). Corcoran responds. *Social Work Research and Abstracts, 22*, 23-24.
- Eggebrecht, D. B., Bautz, M. T., Brenig, M. I. D., Pfingsten, M., & Franz, C. (1989). Psychometric evaluation. In P. M. Camic & F. D. Brown (Eds.), *Assessing chronic pain: A multidisciplinary handbook* (pp. 71-90). New York: Springer-Verlag.
- Eimer, B. N., & Freeman, A. (1998). *Pain management psychotherapy: A practical guide*. New York: John Wiley.
- Gingerich, W. J. (1984). Meta-analysis of applied time-series data. *Journal of Applied Behavioral Science, 20*, 71-79.
- Guyatt, G. H., Keller, J., Jaeschke, R., Rosenbloom, D., Adachi, J. D., & Newhouse, M. T. (1990). The n of 1 randomized controlled trial: Clinical usefulness. *Annals of Internal Medicine, 112*, 293-299.
- Holden, G., Bearison, D., Rode, D., Fishman-Kapiloff, M., Rosenberg, G., & Rosenzweig, J. (2002). The impact of a computer network on pediatric pain and anxiety: A randomized controlled clinical trial. *Social Work in Health Care, 36*, 21-33.
- Holden, G., Bearison, D., Rode, D., Rosenberg, G., & Fishman, M. (1999). Evaluating the effects of a virtual environment (STARBRIGHT World) with hospitalized children. *Research on Social Work Practice, 9*, 365-382.
- Holden, G., Bearison, D., Rode, D., Rosenberg, G., & Fishman, M. (2000). The effects of a computer network on pediatric pain and anxiety. *Journal of Technology in Human Services, 17*(1-3), 27-48.

- Jayarathne, S., Tripodi, T., & Talsma, E. (1988). The comparative analysis and aggregation of single-case data. *Journal of Applied Behavioral Science, 24*, 119-128.
- Litt, M. D., Cooney, N. L., & Morse, P. (1998). Ecological momentary assessment (EMA) with treated alcoholics: Methodological problems and potential solutions. *Health Psychology, 17*, 48-52.
- McGrath, P. A. (1990). *Pain in children: Nature, assessment, and treatment*. New York: Guilford.
- McGrath, P. A., Seifert, C. E., Speechley, K. N., Booth, J. C., Stitt, L., & Gibson, M. C. (1996). A new analogue scale for assessing children's pain: An initial validation study. *Pain, 64*, 435-443.
- Onghena, P. (1992). Randomization tests for extensions and variations of ABAB single-case experimental designs: A rejoinder. *Behavioral Assessment, 14*, 153-171.
- Onghena, P. (1994). *The power of randomization tests for single case designs*. Doctoral dissertation, Catholic University of Leuven, Leuven, Belgium.
- Onghena, P., & Edgington, E. S. (1994). Randomization tests for restricted alternating treatment designs. *Behaviour Research and Therapy, 32*, 783-786.
- Rosnow, R. L., & Rosenthal, R. (1989). Statistical procedures and the justification of knowledge in psychological science. *American Psychologist, 44*, 1276-1284.
- Salzberg, C. L., Strain, P. S., & Baer, D. M. (1987). Meta-analysis for single-subject research: When does it clarify, when does it obscure? *Remedial and Special Education, 8*, 43-48.
- 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.
- Shiffman, S., & Stone, A. (1998). Introduction to the special section: Ecological momentary assessment in health psychology. *Health Psychology, 17*, 3-5.
- Thyer, B. A. (1996). Forty years of progress toward empirical clinical practice? *Social Work Research, 20*, 77-82.
- Tuckman, B. W. (1988). The scaling of mood. *Educational and Psychological Measurement, 48*, 419-427.
- Van Damme, G., & Onghena, P. (1995). *Manual SCRT1.1: Single case randomization tests*. Leuven, Belgium: Catholic University of Leuven, Research Group on Quantitative Methods.
- Varni, J. W., Blount, R. L., Waldron, S. A., & Smith, A. J. (1995). Management of pain and distress. In M. C. Roberts (Ed.), *Handbook of pediatric psychology* (2nd ed., pp. 105-123). New York: Guilford.
- Videka-Sherman, L. (1986). Alternative approaches to aggregating the results of single-subject designs. *Social Work Research and Abstracts, 22*, 22-23.
- Wakefield, J. C., & Kirk, S. A. (1996). Unscientific thinking about scientific practice: Evaluating the scientist-practitioner model. *Social Work Research, 20*, 83-96.
- White, D. M., Rusch, F. R., Kazdin, A. E., & Hartmann, D. P. (1989). Applications of meta-analysis in individual-subject research. *Behavioral Assessment, 11*, 218-296.
- White, O. R. (1987). Some comments concerning "The Quantitative Synthesis of Single-Subject Research." *Remedial and Special Education, 8*, 40-42.
- Witkin, S. L. (1996). If empirical practice is the answer, then what is the question? *Social Work Research, 20*, 69-76.
- Zucker, D. R., Schmid, C. H., McIntosh, M. W., D'Agostino, R. B., Selker, H. P., & Lau, J. (1997). Combining single patient (n-of-1) trials to estimate population treatment effects and to evaluate individual patient responses to treatment. *Journal of Clinical Epidemiology, 50*, 401-410.