Response to Commentaries on A Strengths-Based, Skill-Building, Integrative Approach to Treating Conduct Problems in a 12-Year-Old Boy: Rafael’s Story

PAUL W. CLEMENT a,b,c

a Private practice, South Pasadena, CA
b Correspondence concerning this article should be addressed to Paul W. Clement, 719 Fremont Avenue, South Pasadena, CA 91030.
c I want to thank Daniel Fishman, the editor of this PCSP journal, for inviting me to write a case study on the treatment of a child like “Rafael’s Story” (Clement, 2011), as well as his many suggestions for improving the first draft of the article. Email: PaulWClement@aol.com

ABSTRACT

In this discussion I respond to two commentaries on a case study of my treatment of Rafael, a 12-year-old boy presenting with conduct problems (Clement, 2011b). My response includes shorter consideration of a variety of specific points raised by the commentators. I also elaborate in some detail on one of the points raised concerning my system for measuring treatment outcomes in routine clinical practice. The approach adapts Goal Attainment Scaling, the Global Assessment Scale/Global Assessment of Functioning, and Smith and Glass’s 1977 method of calculating treatment effect sizes (ESs). Finally, I frame my case study as an example of practice-based evidence for the effectiveness of psychotherapy.

Key words: adolescent psychotherapy; birth defects; child psychotherapy; conduct problems; developmental assets; developmental trajectory; positive psychology; practice-based evidence; resilience; stealing; vandalism; clinical case studies; case studies

California requires at least one year of post-doctoral supervised experience before a person can become licensed as a psychologist. During my first year following graduate school I did my clinical work and research in the Children’s Psychiatric Clinic at Harbor/UCLA Medical Center in Torrance, California. The head of the clinic arranged for a psychiatrist on our staff to provide me with the required pre-licensed supervision. The supervising psychiatrist and I had very different theoretical orientations. Unfortunately, we spent our early months of supervision arguing. The process was unrewarding for both of us. In thinking about what had been transpiring in our supervisory sessions, I decided to make a proposal. In our next meeting I reviewed what I perceived had been occurring in our meetings and suggested a change. I proposed that from that day forward I would describe to him what I was doing with each of my therapy cases and my rationale for doing so. Then I wanted him to tell me what he would do with
each of my cases and his rationale for what he would do if the case was his. We would end our discussion at that point without making any attempt to say who was right and who was wrong.

In preparing my response to the comments by Oades-Sese and Kitzie (2011) and by Hawes (2011) on my case study of Rafael, I have applied the same approach as sketched out in the preceding paragraph for the second phase of my pre-licensed supervision. I want to thank them for their thoughtful comments. I will reply selectively to what they have said.

**RESPONSE TO OADES-SESE AND KITZIE’S COMMENTS**

Let me begin by saying that overall, I very much appreciate Oades-Sese and Kitzie's discussion of the literature on resilience in children in their section on "Overcoming Adversity." This discussion is very helpful in placing the resilience themes in Rafael's case in broader context.

One specific point. In their discussion, Oades-Sese and Kitzie (2011) state that Rafael "has recollections that his biological parents placed him in an orphanage at 3 months of age" (p. 400). I don’t believe that he had this kind of recollection, although he did remember being told later in his life that his birth parents placed him in an orphanage when he was an infant and that the mother who raised him adopted him when he was 13 months old. I do agree with their pointing out that Rafael manifested much resilience in the face of various negative circumstances that he faced during his growing years.

I appreciated their brief review of the possible contributions of “Cultural and Ethnic Identity” to resilience in a child and adolescent. Rafael was born on another continent. His racial/ethnic identity was not the same as that of this adoptive mother and siblings. Although some child therapists likely would have explored the meaning to him of his racial/ethnic uniqueness, I did not do so.

In reviewing “Cognitive-Behavioral and Social Skills Interventions,” Drs. Oades-Sese and Kitzie made several suggestions for how I could have improved my treatment of Rafael. Their recommendations all made sense and could have added to my success in treating him. On the other hand, Kazdin (2000) has reported that child and adolescent psychotherapists have used over 500 forms of treatment, and it is not feasible to use even a small proportion of them in a single case. That was certainly true in the present case.

I have little to say regarding Oades-Se se and Kitze's comments about the possible "emotional distance" between Rafael and myself as his therapist. Although Rafael did not become highly skillful in verbalizing his emotions during the time that I treated him, by the time of our final termination his “acting out” behaviors had been replaced by goal-oriented actions that continued to be manifested throughout the eight-year follow-up period. On a different point, I didn’t just minimize myself as a father figure to him; I never conceived myself as such. I saw myself more like a friendly coach, mentor, and “rooting section” for his life in the present and near future. Although he did not set goals for himself to be pursued through psychotherapy, he did articulate many personal goals during our sessions. He continued to set and pursue his own goals during the eight years of follow-up and continues to do so into the present.
The suggestions by Oades-Sese and Kitzie on the role of “Assessment of Behaviors, Quality Of Life, and Coping Strategies” in Rafael’s case stand on their own. These commentators don’t call for a response from me. Similarly, I found their proposal to use “Narrative Therapy” an interesting idea. I think that both Rafael and I would have enjoyed including that approach.

Oades-Sese and Kitzie raise a question about “Possible Undiagnosed ADHD.” During the course of my 46-year career of diagnosing and treating children and adolescents, the most frequent diagnosis made by me has been attention-deficit/hyperactivity disorder. The second most frequent diagnosis has been oppositional defiant disorder. Rafael presented with some symptoms that were consistent with both of these diagnoses, but in my assessment his symptoms did not rise to the number or severity required for either diagnosis.

In my view, the comments by Oades-Sese and Kitzie appearing in the section labeled “Outcome Measurement” call for the most detailed response on my part. They say, “Clement presents no evidence of validity or reliability in his discussion of this instrument [Childhood Problems Checklist]. In addition, Rafael’s mother was given a photocopy with her previous ratings on it to complete during the latter two sessions. This calls into question the validity of the results” (p. 407).

For the Childhood Problems Checklist, the available reliability data are based on inter-rater agreement when two parents independently fill out the checklist. The mean Pearson $r$ has been 0.58 ($SD = 0.15$) (95% CI: 0.53-0.64). This mean happens to be identical to that reported by Achenbach (1991) for the Child Behavior Checklist/4-18. This latter instrument is probably the most widely used instrument to measure treatment outcomes in controlled treatment outcome studies with children and adolescents.

The Scale of Functioning (SOF) used with the Childhood Problems Checklist is a truncated form of the Global Assessment Scale (GAS) (Endicott, Spitzer, Fleiss, & Cohen, 1976) and of the Children’s Global Assessment Scale (CGAS) (Shaffer, Gould, Brasic, Ambrosini, Fisher, Bird, & Aluwahlia, 1983). For both the GAS and CGAS, users are instructed to select the lowest level of functioning in the patient at the time of the intake evaluation. In adapting this instruction to the Childhood Problems Checklist I use the following simple formula: $[(mean \ SOF \ score \ at \ intake) \ - \ (SD \ of \ the \ SOF \ scores \ at \ intake)]$. This produces an estimated Global Assessment of Functioning (GAF) score. In my private practice the mean estimated GAF score at intake using the Childhood Problems Checklist has been 42.56 ($SD = 12.17$) (95% CI: 40.39-44.74). Shaffer et al. (1983) claimed, “Scores above 70 on the CGAS are designated as indicating normal function” (p. 1228); hence, the estimated GAF score seems to be a clear indicator of a low level of functioning at intake and suggesting the need of treatment.

The formula for determining treatment effect size (ES) is as follows: $[(mean \ SOF \ score \ at \ termination) \ - \ (mean \ SOF \ score \ at \ intake)]/(SD \ of \ SOF \ scores \ at \ intake)$. Using the Childhood Problems Checklist the mean ES across all my cases at termination has been 1.83 ($SD = 1.42$)(95% CI: 1.58-2.08). Using the SOF and the checklist seems to provide a sensitive measure of change during a course of treatment. For children seen in my practice, ESs at termination have ranged from -0.39 to 7.43, with 3.60% of all ESs have been negative.
When Rafael returned for the second course of treatment, he was too old for the Childhood Problems Checklist. I had him and his mother use the SOF to rate him independently on four areas of concern at intake and again at termination. The correlation between their ratings at intake was 0.58 across the four areas. The correlation between their ratings at termination was 1.00. The correlation between their ratings when combining the eight ratings at intake and termination was 0.94. As indicated in Table 5 of the main article (Clement, 2011b), the ES based on the input of Rafael and his mother was 3.02.

Now I would like to review the evolution of the methods that I used in measuring outcomes during my two courses of treatment with Rafael. Very early in my career I concluded that the traditional psychological tests that I had learned to use during graduate school were not useful for tracking treatment outcomes in routine clinical practice. In my treatment outcome research (e.g., Clement & Milne, 1967; Clement, Fazzone, & Goldstein, 1970) I leaned heavily on trained research assistants to observe and record the behavior of participants in therapy, but I could not employ trained research assistants to observe patients in my private practice. For some private cases a parent was able to gather observational data. For example, early in my career I treated a young boy for sleep walking, and his mother kept written records of his sleep-walking episodes (Clement, 1970). Unfortunately, for many or most cases direct observation of behavior was not a practical option.

Shortly after its first published description I discovered Goal Attainment Scaling (GAS) (Kiresuk & Sherman, 1968). Although the creators of GAS designed it for evaluating comprehensive community mental health programs, others have applied it to a wide range of problems and settings including the following: brain injury, cerebral palsy, chronic pain, communication disorders, dementia, developmentally disabled, diabetes, executive coaching, exercise programs, inpatient child psychiatric services, learning disabilities, nursing, occupational performance, occupational therapy, psychotherapy outcomes, practicum training, residential treatment, sensory integration therapy, social service agencies, special education, summer camps, therapeutic day camps, and violent mentally disordered offenders. I applied GAS to selective cases in my private practice and taught the method to my graduate students during the late 1960s and throughout the 1970s.

GAS requires setting multiple goals for a given program or case. Each goal contains 5 levels:

(-2) much worse than expected outcome,
(-1) worse than expected outcome,
(0) the expected outcome,
(+1) better than expected outcome, and
(+2) much better than expected outcome.

The type and number of goals is individually selected. Each of the five levels for a given goal is also individualized and has specific boundaries. These facts meant that GAS fit well with the idiographic approach of routine clinical practice, but it was labor-intensive. I continued to look for a method that could be used with all of my patients and that did not require so much time to employ.
Starting in the early 1970s my research and that of most of my doctoral students employed single-subject research designs (cf., Kazdin, 2011). These designs have many similarities to treating individual cases in routine clinical practice. But the investigators who tended to use single-subject designs tended to insist on using visual inspection of graphed data to draw conclusions and to reject statistical description and analysis (e.g., Parsonson & Baer, 1992). In 1977 Smith and Glass introduced meta-analysis as a method for performing quantitative reviews of research on psychotherapy. I decided to adapt their method to my single-subject research (e.g., Clement, Anderson, Arnold, Butman, Fantuzzo, & Mays, 1978).

In late 1988 I began to perform a quantitative analysis of all cases that I had seen in my private practice. I was concerned by what I found. There was great variability in how I had tracked outcomes across my cases. I continued my review until I had covered all cases seen through mid 1992 and published the results in *Professional Psychology: Research and Practice* (Clement, 1994). I wanted to develop better methods for measuring treatment outcomes in routine clinical work.

A year before I began the analysis described in the previous paragraph, *DSM-III-R* appeared (American Psychiatric Association, 1987). This revision of the *DSM* included Axis V: the Global Assessment of Functioning Scale. The GAF Scale was based on the Global Assessment Scale (Endicott et al., 1976) mentioned earlier as well as on the CGAS (Shaffer et al., 1983). I decided to blend Goal Attainment Scaling with the GAF Scale. Doing so allowed me to develop a unique list of problems, issues, or concerns for each patient and to create a Scale of Functioning (SOF) for each target. Table 1 provides an example of such a list of problems for an 8-year-old girl diagnosed with oppositional defiant disorder. Table 2 is an example of a tailor-made scale corresponding to the fourth entry in Table 1. There were seven of these tailor-made scales; one for each problem in Table 1. The girl's mother and I worked together to create these seven scales. Then the mother identified at what level her daughter was functioning at intake. At termination she again selected the level of functioning for each of the seven problems. This approach is consistent with the procedures of Goal Attainment Scaling used in program evaluation for over 40 years.

The process described in the preceding paragraph took much time and cooperation by the mother. Not all parents or patients were willing or able to invest the time and effort required to create SOFs for each of their concerns. I looked for a more efficient, less demanding approach. I realized that a modest number of issues were identified as concerns for a majority of my patients. This conclusion led me to create a set of checklists that included the most common intake problems for children, adolescents, adults, and couples. I included the SOF as part of each checklist in order to identify the severity of each problem, GAF score at intake, and eventual treatment outcome expressed as an ES. The Childhood Problems Checklist (Clement, 1999) was one of them.

When a parent, adolescent, or adult patient uses the SOF to indicate how well the patient is doing at a particular point in time, I do not have a way to determine how much of the variance in their ratings is due to their desire to impress me with how bad they perceive their situation at intake or to make me feel good during the course of treatment or at termination. I am clear that
parents decide to bring their child to a therapist, and the same thing is usually true when an adolescent is the patient. Parents also play a key role in deciding when to end treatment. Given these facts, involving parents in evaluating treatment outcomes seems essential.

The biggest problem I have encountered in employing my current approach to measuring treatment outcomes is when patients drift away without having a termination interview. When that has happened I have had much difficulty obtaining SOF scores at a later time. Often I have faced the task of estimating outcome from all of the evidence contained within a given patient’s chart, and that is what I started doing in late 1988 (cf. Clement, 1994). I assign each completed treatment case a Global Estimate of Outcome (GEO) score as follows:

1 = Much Worse (>49% worse than at intake or ES = -1.50 or less),
2 = Worse (11-49% worse than at intake or ES = -0.51 to -1.49),
3 = No Change (within 10% of intake or ES = -0.50 to +0.50),
4 = Improved (11-49% better than at intake or ES = +0.51 to +1.49), and
5 = Much Improved (>49% better than at intake or ES = +1.50 or greater).

In analyzing my treatment outcomes during my first 42 years of private practice for patients of all ages, I compared GEO scores based on my review of the records when I didn’t have ES scores based on patient or parent ratings with GEO scores that were based on ES scores (Clement, 2011a). The mean GEO score based on my judgment was 3.80 ($n = 991$, $SD = 0.80$), and it was 4.40 ($n = 582$, $SD = 0.74$) when based on ESs derived from patient or parent ratings ($F = 215.50$, $df = 1$ & $1,571$, $p = 0.0000$). For whatever reasons, parents and adolescent patients have rated our treatment outcomes more favorably than I have.

**RESPONSE TO HAWES’ COMMENTS**

I appreciated Hawes’ (2011) review of research published in the past decade on developmental trajectories of antisocial behavior. Hawes added much to what I covered in telling Rafael’s story. His discussion of sub-typing risk pathways to antisocial behavior was also very helpful. Although adolescents manifesting callous-unemotional traits have not constituted a significant portion of my referrals, those who have presented with reduced guilt and empathy have been among my more troubling cases.

I have never used the Antisocial Process Screening Device (Frick & Hare, 2002). Given Hawes’ description of this instrument, I do not believe that Rafael would have received high scores that predicted serious antisocial behavior. I would have liked to know, however, the degree to which my clinical judgment and the scores obtained on this instrument would have been in agreement.

Had I concluded that Rafael met the criteria for conduct disorder or for oppositional defiant disorder, I would certainly have worked within a family therapy format. In fact, I have used family therapy with 63.41% of my child and adolescent cases. I have used individual therapy with only 35.37% of my cases. For the remaining 1.22% I have used tailor-made group therapy.
Regarding the duration of my treatment of Rafael, his lasted much longer than most of my child and adolescent cases. In my career the mean number of sessions per case has been 16.16 (SD = 20.73) (95% CI: 14.58—17.73) and the median has been 10.

Toward the end of his comments Dr. Hawes proposed, “a session structure that permits the therapist to build relationships with the adolescent and his/her parents separately (e.g., dividing the session to accommodate one-on-one time with respective subsystems)” (p. 418). I fully agree with this suggestion. For a large proportion of my cases treated via “family therapy,” I have split the sessions in half. I spend half of the session with the parent(s) and half with the youth. This strategy has worked particularly well within the structure of a private practice.

FINAL RESPONSE

During my years as a student in university and graduate school (1957-1965), there was little empirical evidence that psychotherapy made a positive difference (Eysenck, 1952; Levitt, 1957). Like many psychologists of my vintage, I was motivated to find evidence for the benefits of psychotherapy. At the beginning of my career I did empirical research on psychotherapy with children using between-groups experimental designs. By the early 1970s, however, I shifted to single-subject experimental designs.

In both my research and my private practice I was interested in measuring treatment outcomes empirically. I was encouraged when Meltzoff and Kornreich (1970) published their comprehensive review of research on psychotherapy. They concluded that there was substantial evidence that psychotherapy is more beneficial than no psychotherapy. Seven years later Smith and Glass (1977) employed “meta-analysis” to review controlled psychotherapy outcome studies quantitatively. They were able to state by how much therapy was more beneficial than no treatment. Three years later, Smith, Glass, and Miller (1980) expanded the earlier review into a book-length analysis. Although some psychologists criticized their methods of meta-analysis, many other reviewers have used and adapted their methods to measuring treatment effects both within various areas of psychology as well as in other disciplines.

During the 1980s and 1990s reviewers used meta-analysis to determine the efficacy of various psychological treatments with many clinical populations. Efficacy studies usually employed randomized clinical/controlled trials (RCTs). This work culminated in Division 12 of APA listing criteria to identify empirically-validated psychological treatments (Task Force, 1995) and establishing an initial list of such treatments. In response to protests against the term, validated, Division 12 substituted a softer term and began referring to empirically supported treatments (Chambless et al., 1996). Nevertheless, for the next decade protests against such lists and frequent debates at the annual conventions of APA followed.

The protestors were eventually successful in getting APA to adopt a policy statement on evidence-based practice (APA Presidential Task Force, 2006). This policy statement helped to set the stage for balancing efficacy findings from RCTs with effectiveness data from practice-based evidence. Although researchers have been slow to answer Seligman’s (1995) call for effectiveness studies, there is growing interest in many disciplines to incorporate practice-based
evidence as an important part of the facts underlying the evidence-base for their profession (e.g., Barkham, Hardy, & Mellor-Clark, 2010; Boba, 2010; Green, 2008; Grissom & Lyons, 2006; Mapp, Boutté-Queen, Erich, & Taylor, 2008). Pragmatic Case Studies in Psychotherapy is an important journal for presenting practice-based evidence, and I have welcomed the opportunity to add Rafael’s Story to the developing database of systematic case studies the journal publishes.

REFERENCES


Clement, P.W., & Milne, D.C. (1967). Group play therapy and tangible reinforcers used to modify the behavior of eight-year-old boys. Behaviour Research and Therapy, 5, 301-312.


Table 1

Sample Master List of Problems, Issues, or Concerns for an 8-Year-Old Girl Diagnosed with Oppositional Defiant Disorder

<table>
<thead>
<tr>
<th>Description</th>
<th>Pre-Treatment SOF</th>
<th>Post-Treatment SOF</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mother-daughter arguments and fights</td>
<td>3</td>
<td>9</td>
</tr>
<tr>
<td>Daughter complies with time-out instructions</td>
<td>3</td>
<td>9</td>
</tr>
<tr>
<td>Unhappy and bored in school</td>
<td>4</td>
<td>9</td>
</tr>
<tr>
<td>Tantrums</td>
<td>2</td>
<td>9</td>
</tr>
<tr>
<td>Sleeping in her mother’s bed since father left 13 months ago</td>
<td>3</td>
<td>9</td>
</tr>
<tr>
<td>Does not accept the end of her parents’ marriage</td>
<td>5</td>
<td>9</td>
</tr>
<tr>
<td>Oppositional, defiant, noncompliant</td>
<td>4</td>
<td>8</td>
</tr>
</tbody>
</table>

Mean SOF = 3.43 8.86
SD of SOF Scores = 0.98 0.38
ES = 5.54
Table 2

Sample Scale of Functioning for “Tantrums”

<table>
<thead>
<tr>
<th>Description</th>
<th>Score</th>
<th>Definition of Each Level</th>
<th>Date of Evaluation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Good Functioning</td>
<td>9</td>
<td>No tantrums in past week</td>
<td>09/14</td>
</tr>
<tr>
<td>Slight Problems</td>
<td>8</td>
<td>&lt; 1 tantrum per week</td>
<td></td>
</tr>
<tr>
<td>Some Problems</td>
<td>7</td>
<td>1 tantrum per week</td>
<td></td>
</tr>
<tr>
<td>Moderate Difficulty</td>
<td>6</td>
<td>2 tantrums per week</td>
<td></td>
</tr>
<tr>
<td>Serious Problems</td>
<td>5</td>
<td>3 tantrums per week</td>
<td></td>
</tr>
<tr>
<td>Major Impairment</td>
<td>4</td>
<td>4 tantrums per week</td>
<td></td>
</tr>
<tr>
<td>Inability to Function</td>
<td>3</td>
<td>5 tantrums per week</td>
<td></td>
</tr>
<tr>
<td>Some Danger of Hurting Self or Others</td>
<td>2</td>
<td>6 or more tantrums per week</td>
<td>02/17</td>
</tr>
<tr>
<td>In Persistent Danger of Hurting Self or Others</td>
<td>1</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>