This paper outlines the purposes, professional obligations, and key components to consider when providing effective evaluation in psychotherapy supervision. An overview of various methods for gathering supervision data for evaluation purposes is provided including self-reporting; process notes; video and audiotaping; live observation; co-therapy; and practices of reflectivity. A review is provided of research findings related to evaluation in supervision. The paper describes the creation of a Supervision Outcomes by Method Evaluation Matrix to help supervisors identify the kinds of outcomes they are trying to facilitate and the methods they may want to employ to evaluate those outcomes. A list of references to articles on evaluation in supervision and on instruments and formats for evaluation in supervision are provided. (JDM)
Evaluating Effective Supervision

Vaughn E. Worthen, Ph.D.
Brigham Young University
M. Kirk Dougher, Ph.D.
Brigham Young University


Correspondence concerning this material should be addressed to Vaughn E. Worthen, Counseling and Career Center, Brigham Young University, 2570 WSC, Provo, UT 84602 or M. Kirk Dougher, Counseling and Career Center, Brigham Young University, 1500 WSC, Provo, UT 84602. Electronic mail may be sent to:

vaughn_worthen@byu.edu
kirk_dougher@byu.edu
Abstract

This paper outlines the purposes, professional obligation, and important components necessary for consideration in providing effective evaluation in psychotherapy supervision. We give a brief overview of various methods for gathering supervision data for evaluation purposes, including: self-report, process notes, video and audiotaping, live observation, co-therapy, and practices of reflectivity. Next we present a concise review of research findings related to evaluation in supervision. These issues include such items as: the role of the relationship, identifying, developing and communicating evaluative criteria to the supervisees, instrumentation in evaluation, the role of goal setting and feedback, understanding the context of supervisee development in evaluation, the role of evaluation in promoting self-reflective practice, resistance to evaluation in supervision, and biases that can operate in evaluation. We also created an "Supervision Outcomes by Method Evaluation Matrix" to help supervisors identify the kinds of outcomes they are trying to facilitate and the methods they may want to employ to evaluate those outcomes. Finally, we provide a list of references to articles on evaluation in supervision and another list of references to instruments and formats for evaluation in supervision.
Evaluating Effective Supervision

Vaughn E. Worthen, Ph.D. & M. Kirk Dougher, Ph.D.

Introduction

Although it may be described in different terms from various theoretical orientations, most supervisors would likely agree, the goal of supervision is to direct the education of therapists-in-training in order to increase their repertoire of effective skills (Williams, 1988), thus enhancing outcomes with clients. With supervision aiming to increase a trainee's repertoire of techniques and personal abilities, supervisors shoulder the burden of knowing what skills a therapist-in-training requires, how to develop this repertoire and when they have reached the goals developed for supervision.

As Worthington (1987) suggests, supervisors have little training in how to supervise effectively. Moreover, supervisors not only lack training in how to supervise but also have limited availability to methods of recognizing improvements in supervisee skills.

We will attempt to identify relevant findings that are designed to augment discussion and knowledge about the process of supervision evaluation. This review is not designed to be a comprehensive work, rather an overview. It represents some references and ideas that those faced with the daunting task of evaluating supervision, and the improvements therein, might find useful.

Purpose of Supervision Evaluation: To monitor client welfare and improve client outcomes as well as foster the development of new professionals through the enhancement of supervisee skill, self-efficacy, professional identity, sensitivity and ability to work with diverse populations, ethical and legal compliance, etc.
Evaluation can be divided into two main tasks: Goal-setting (the setting of criteria) and Feedback (how well the criteria are being met) (Lehrman-Waterman & Ladany, in press).

**Ethical Obligation to Evaluate:**


**Evaluation Components** (Ellis & Ladany, 1997):

- Mode of therapy (individual, group, couples, family)
- Domain of trainee behaviors (therapy or supervision)
- Competence area (techniques, conceptualization, ethical behavior, specific skills, supervision behaviors, etc.)
- Method of evaluation (self-report, case notes, videotape, live, etc.)
- Proportion of caseload (are you evaluating all clients, a subgroup [multicultural], or a single client)
- Segment of experience (the complete duration of therapy, a specific session, a segment of a session)
- Sequence of training (early in training, middle of treatment course, or end)
- Evaluator (supervisor, supervisee, client, peers, objective rater)
- Level of proficiency (beginner, more advanced)
- Reliability (is the evaluation instrument you are using reliable or if qualitative is it trustworthy, etc.)
- Validity (is it measuring what you believe it should be measuring)
• Format (quantitative/qualitative)

General Methods of Data Gathering for Evaluation:

• **Self-Report:** (Supervisee verbal reports without other supporting documentation of therapeutic activities.) This is perhaps the most commonly used form of evaluation (Borders, Cashwell, & Rotter, 1995; Romans, Boswell, Carlozzi, & Ferguson, 1995) and one estimate has it occurring 90% of the time (Ladany & Lehrman-Waterman, 1997). Formative evaluation efforts specifically rely heavily on self-report. Self-report is subject to distortion, self-selection, and frequently is not a very rich data source. Holloway (1985) indicated that self-report was particularly inadequate with novice trainees. In one study (Rogers & McDonald, 1995), more direct evaluation methods led supervisors to assess their supervisees as being less prepared than when relying more on self-report.

• **Process Notes:** (Supervisees written summaries of their work with clients, including session notes). This method can be used as an adjunct to other methods of evaluation. Evaluating process notes allows the supervisor to see if record keeping is appropriate including documenting legal and ethical issues, actions, and obligations. It may also give a supervisor another perspective on the supervisees' conceptual ability including diagnosis, treatment plan, interventions, etc. Many of the same concerns occur with this method as the previous method of self-report.

• **Video and audiotaping:** Strengths include allowing the supervisor to view or hear direct behavior or verbal responses from the trainee and client, providing a rich source of data. On the other hand this rich data may be overwhelming if observations are not focused and goal directed. Many authors have generated recommendations about using tapes
(Aveline, 1997; Brandell, 1992; Breunlin, Karrer, McGuire, & Cimmarusti, 1988; Cashwell, Looby, & Housely, 1997; Kagan, 1980). Ladany and Lehrman-Waterman (1997) indicate that this medium is used less than 60% of the time. There has been a general concern that the intrusive nature of taping, especially videotaping would lead to performance anxiety in supervisees and decreased performance. Ellis, Beck, & Krengel (1998) found in a study of the effects of increased awareness of self (videotaped sessions) that “...audio-videotaping neither significantly increases counselor trainees’ anxiety, nor decreases their performance” (p. 9), thus, they concluded, “A growing body of evidence appears to suggest that the negative effects of audio-videotaping supervisees is transitory or negligible” (p.9).

- **Live Observation:** This is sometimes accomplished by just being an observer, viewing a session from behind a one-way mirror. At other times live observation can also incorporate the supervisor as an active participant. Some examples of active participation include the “bug-in-the-ear,” telephoning into a session, or consultation breaks in a session. A recently developed technique also provides more liberty from rule-governed behavior. This method includes a computer “monitor-in-the-corner” (Follette & Callaghan, 1995). Gould and Bradley (in press) summarize six advantages to live observation: 1) increased likelihood of positive counseling outcomes, 2) supervisees learn more efficiently through this method, 3) clients are more directly protected, 4) supervisees can work with more challenging clients, 5) supervisees tend to risk more, and 6) the supervisory relationship may be enhanced. They also identify some disadvantages including, the time and resources needed to tape sessions and then review tapes, discomfort of clients and supervisees, may effect performance since supervisee may
attempt to impress supervisor or they may be inhibited and not attempt new or creative interventions.

- **Co Therapy**: An often-overlooked method of gathering information about the skills and progress of therapists-in-training is to engage in co-therapy with your supervisee. This method provides real-time data on the therapeutic interplay between the student-therapist and the client. This method does come with a number of caveats that include time commitment and deferring to the supervisor during the session. However, if these issues are addressed, co-therapy can have tremendous utility through modeling.

- **Practice and Reflectivity**: This method is comprised of several procedures that can occur during the supervision. First, role-playing. This allows for the supervisor to depict various hypothesized client behaviors. It demands the student to react to the dynamic nature of the therapy session rather than relying on the supervisor to script upcoming sessions. Second, alternative hypothesis generation. In this procedure supervisors suggest alternate data that may become available to the therapist-in-training. Rather, than display the behavior they might consider, as in the role-play, students are encouraged to develop case conceptualization and treatment planning based on newly acquired information. In both role-playing and alternative hypothesis generation student-therapists exhibit behaviors that are central to those being evaluated and developed. And third, engaging in reflective analysis. Here supervisors encourage supervisees to use self-reflective practice to enhance their repertoire development (Neufeldt, 1999; Schon, 1983).
General findings:

- As is true in effective supervision in general the supervision relationship plays a significant role in providing effective evaluation. The supervisory relationship is critical in creating an environment where effective evaluation can occur and be received as a learning experience for supervisees (Talen & Shindler, 1993). Effective goal setting and feedback are associated with stronger supervisory working alliance (Lehrman-Waterman & Ladany, in press). The challenge is not to allow the supervisory relationship to unduly influence the evaluations so that evaluations are either too positive or too negative based on the relationship and not fully informed using competence criteria. Lazar and Mosek (1993) found some evidence that the working relationship did influence evaluations so that competency evaluations were biased by the relationship. Also Dodenhoff (1981) found that interpersonal attraction between the supervisee and supervisor predicted how positive the evaluation would be. Liked supervisees also receive more effort and support from supervisors and more positive evaluations than disliked supervisees (Turban, Jones, & Rozelle, 1990). Thus, the impact of the supervisory relationship is significant when it comes to evaluation. Care should be taken to foster an effective working relationship, which includes the establishment of clear criteria and methods for supervision evaluation so as to promote objectivity and accuracy in evaluation rather than relying on a global sense of liking or disliking.

- Supervisees are primarily evaluated using qualitative procedures (Norcross & Stevenson, 1984).

- Supervisors may not be adequately performing their evaluative roles (Ladany, Lehrman-Waterman, Molinaro, & Wolgast, 1999).
Few theorists have addressed how supervisors change as they gain experience. Bernard (1979, 1981, and 1982) described how supervisors could be trained to increase the awareness of options during supervision. She outlined a training method equally applicable to beginning and experienced supervisors. She reported supervisors often become aware of a discrepancy between their intention and performance. The most common discrepancy is supervisors think they are using a counselor role, but in actuality, they engage in a teacher role.

Alternately, Alonzo (1983) proposed that supervisors continue to change throughout their professional life cycle. She suggested change from novice to mid-career to late-career concerns. At each stage, it was suggested, supervisors wrestle with three issues: self and identity, the relationship between therapist and supervisor and the administrative duties they have. The supervisors within this model are said to resolve these issues differently, because the demands of life and professions differ at each stage of their career.

Another view is the components of supervision are similar to therapy and having skills as a therapist equates to having the necessary skills to be a supervisor (Rubin, 1989). If this assumption were true, then no further (or only limited) training beyond therapy would be necessary to become a good supervisor. In addition, it would not be necessary to conduct research on a subject the profession implicitly agrees is being handled well. The fact that supervision is being researched implies considerable disagreement with this assumption.

There are differences in skillfulness in supervision among supervisors (Worthington, 1987). Several investigations into the effectiveness of supervision have found a variety of levels of competence and impact (Cross & Brown, 1983; Heppner & Handley, 1981; Worthington 1984; Zucker & Worthington, 1986). Moreover, their competence level, as
rated by their supervisees, does not seem to improve with experience (Marikis, Russell & Dell, 1985).

- Goals should be clear, specific and measurable (Lehrman & Ladany, in press).
- Criteria and methods for summative evaluation should be explicitly discussed early in supervision (McCarthy, et al., 1995; Lehrman-Waterman & Ladany, in press). Ongoing evaluation will promote greater growth and avoid surprises at the end of the supervisory experiences that could flavor the experience in a negative fashion. Most suggest at least two points of summative evaluation, mid-term and at the end, but others could be included.
- Formative Evaluations occur more regularly and systematically than Summative evaluations.
- There are very few reliable and valid instruments for assessing supervision (Ellis & Ladany, 1997).
- Most evaluation is by the supervisor, often based on very few data points (e.g., such as only one method like video-tapes, process notes, multicultural competencies, etc.).
- Effective evaluation that is conducted from a supportive stance can strengthen the supervisory alliance (Lehrman-Waterman & Ladany, in press; Patton & Kivlighan, 1997) and decrease role conflict and ambiguity in supervisees (Ladany & Friedlander, 1995). One study found that when supervisors liked their supervisees they evaluated them more favorably (Turban, Jones, & Rozelle, 1990).
- Effective goal setting and feedback can lead supervisees to view their supervisors as having contributed to their self-efficacy (Lehrman-Waterman & Ladany, in press).
• Effective goal setting and feedback leads to increased satisfaction with supervision for supervisees (Kadushin, 1992; Lehrman-Waterman & Ladany, in press).

• Evaluation should take into account the experience or developmental level of the supervisee (Stoltenberg, McNeill, & Delworth, 1998), but regardless of experience level of the supervisee goal setting and feedback continue to remain important (Lehrman-Waterman & Ladany, in press).

• Evaluation efforts should ultimately lead to self-reflective practice (Neufeldt, 1999; Schon, 1983).

• Collaborative goal setting and evaluation is desirable (Talen & Schindler, 1993).

• Feedback should be timely (Freeman, 1985; Lehrman-Waterman & Ladany, in press), generally focused on specific behaviors (Borders & Leddick, 1987; Freeman, 1985), based on direct observation (Lehrman-Waterman & Ladany, in press), given in a manner that balances constructive feedback with acknowledging strengths (Halgin & Murphy, 1995; Lehrman-Waterman & Ladany, in press; Stoltenberg, McNeill, & Delworth, 1998; Henderson, Cawyer, & Watkins, 1999), conducted with an appreciation for the sensed vulnerability of supervisees (Henderson, Cawyer, & Watkins, 1999) and include both verbal and written feedback.

• Supervisors may resist evaluation for the following reasons: never been trained adequately in evaluation (Borders & Leddick, 1987), difficulty in switching roles from a nurturing therapist to gatekeeper and supervisor that may need to give constructive feedback that challenges or points out weaknesses, personal avoidance of perceived conflict (Robiner, Fuhrman, & Ristvedt, 1993), and the lack of good methods and instruments for conducting evaluations. Also supervisors are hesitant about evaluating...
supervisees unless they have been working together for a period of time (Fried, Tiegs, & Bellamy, 1992). Related to this finding, Blodgett, Schmidt, and Scudder (1987) found that supervisors rate the same trainees different according to how long they have known the trainee; generally favoring the positive side for the longer they have known the trainee.

- Supervisors can be affected by various biases. Halo effects, leniency, strictness, or central tendency biases may occur when evaluations are globally positive because supervisors like their supervisees, lenient when there are vague criteria for evaluation, strict when expectations are too high or experience level is not taken into consideration, or when supervisors avoid making distinctions in criteria either because of loosely defined rating measures or desire to avoid being questioned regarding their evaluations (Ellis & Ladany, 1997; Carey, Williams, & Wells, 1988; Robiner, et al., 1993). Also supervisors must be careful not to allow first impressions to overly influence their evaluations, but attempt to stick to evaluating defined criteria. This can be a challenge, since Bernard (1982) found that supervisors might rate the same trainee differently on the same criteria depending upon the importance the supervisor places on the various criteria. One of the influences on deciding what are important criteria for supervisory focus is supervisor theoretical orientation. In relation to this, Goodyear and Robyak (1982) found that theoretical orientation affected the focus in supervision.

- Evaluation procedures and criteria should be discussed early in the supervision relationship as part of informed consent and role induction procedures and can include the developing of a supervision contract (McCarthey, Sugden, Koker, Lamendola, Maurer, & Renninger, 1995). One of the effects of unclear evaluation standards or methods is
that supervisees may not disclose some information because of their concerns over evaluation. In a study of nondisclosures of supervisees, evaluation concerns was listed as one of the top 4 reasons (44% of supervisees) for not disclosing in supervision (Ladany, Hill, Corbett, & Nutt, 1996).

- Training supervisees in self-evaluation skills is important (Borders, et al., 1991; Munson, 1983). Dowling (1984) found that trainees were fairly accurate self-evaluators and also good peer evaluators. Self-evaluation can be accomplished in many ways; such as having trainees periodically examine a tape in some detail for intent, response patterns, use of silence, etc. Self-evaluation is one of the important skills that lead to self-reflective practice (Neufeldt, 1999; Schon, 1987).

- Ratings of poor performance should not be a surprise to the supervisee. Ongoing evaluation is important. Goals, feedback, and summative evaluation materials should be documented.

- Concern about evaluation was one of the main reasons for nondisclosure in a group of interns, who generally reported their supervisory relationship to be good on the whole (Hess, S., Schultz, J., Knox, S., Sloan, L., Brandt, S., Kelley, F., & Hoffman, M. A., 1999).

- “Students viewed the process of evaluation to be central to effective supervision” (Henderson, Sawyer, & Watkins, 1999).
**Supervision Outcomes by Method Evaluation Matrix**

<p>| Evaluation Data Gathering Methods → (source) → | SUMMATIVE EVALUATION (Did supervisee meet acceptable criterion standard) | A. Supervisee self-evaluation | B. Reports by supervisors | C. Client input | D. Client Outcome measures | E. Pre-Post measures | F. Formal instrumentation (i.e., SWAI, SSI, RCRU, etc.) | G. Evaluation of written materials (i.e., notes, assessment reports, etc.) | H. Expert/other observations | FORMATIVE EVALUATION (continuous feedback) | I. Audio-video taped observations | J. Live supervision | K. Supervisee self-report | L. Process notes | M. Peer feedback through group supervision | N. Self-evaluation/self-feedback | O. Co-therapy |
| <strong>Supervision Outcomes</strong> | ↓ | ↓ | ↓ | ↓ | ↓ | ↓ | ↓ | ↓ | ↓ | ↓ | ↓ | ↓ | ↓ | ↓ | ↓ | ↓ | ↓ | ↓ | ↓ |
| ↓ | ↓ | ↓ | ↓ | ↓ | ↓ | ↓ | ↓ | ↓ | ↓ | ↓ | ↓ | ↓ | ↓ | ↓ | ↓ | ↓ | ↓ | ↓ | ↓ |
| <strong>PROFESSIONAL DEVELOPMENT</strong> | | | | | | | | | | | | | | | | | | | | | |
| • 1. Confidence (Self-efficacy) | | | | | | | | | | | | | | | | | | | | | |
| • 2. Cognitive Complexity (can tolerate ambiguity, refined clinical judgment) | | | | | | | | | | | | | | | | | | | | | |
| • 3. Professional Identity &amp; Generativity (i.e., commitment to profession, willingness to supervise/mentor, accepts personal responsibility) | | | | | | | | | | | | | | | | | | | | | |
| • 4. Continuous Learning Orientation (i.e., open, flexible, cooperative, accepts feedback) | | | | | | | | | | | | | | | | | | | | | |
| • 5. Ethical and Legal Professional Practice | | | | | | | | | | | | | | | | | | | | | |
| • 6. Others... | | | | | | | | | | | | | | | | | | | | | |
| <strong>SKILLS IMPROVEMENT (COMPETENCE)</strong> | | | | | | | | | | | | | | | | | | | | | |
| • 7. Establishes an | | | | | | | | | | | | | | | | | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>15</td>
<td>effective therapeutic working alliance</td>
</tr>
<tr>
<td>8</td>
<td>Conceptual and diagnostic skills</td>
</tr>
<tr>
<td>9</td>
<td>Intervention Skills (i.e., following EST protocols, meeting standard of care)</td>
</tr>
<tr>
<td>10</td>
<td>Documentation and record keeping</td>
</tr>
<tr>
<td>11</td>
<td>Termination and referral skills</td>
</tr>
<tr>
<td>12</td>
<td>Crisis intervention skills</td>
</tr>
<tr>
<td>13</td>
<td>Assessment and report writing skills</td>
</tr>
<tr>
<td>14</td>
<td>Group leader skills</td>
</tr>
<tr>
<td>15</td>
<td>Multicultural and diversity awareness and sensitivity</td>
</tr>
<tr>
<td>16</td>
<td>Others...</td>
</tr>
<tr>
<td>17</td>
<td>POSITIVE CLIENT OUTCOMES</td>
</tr>
<tr>
<td>18</td>
<td>INCREASED SELF-AWARENESS (aware of impact on others, knowledge of strengths and weaknesses)</td>
</tr>
<tr>
<td>19</td>
<td>SUPERVISEE SATISFACTION</td>
</tr>
<tr>
<td>20</td>
<td>INCREASED KNOWLEDGE OF CLINICAL LITERATURE</td>
</tr>
<tr>
<td>21</td>
<td>OTHERS...</td>
</tr>
</tbody>
</table>
References

Evaluating Effective Supervision

Vaughn E. Worthen, Ph.D. & Kirk Dougher, Ph.D., 2000


References

Evaluating Effective Supervision Instruments and Formats

Vaughn E. Worthen, Ph.D. & Kirk Dougher, Ph.D., 2000


Title: Evaluating Effective Supervision

Author(s): Vaughn E. Worthen, Ph.D. & M. Kirk Dougher, Ph.D.

Corporate Source: Brigham Young University

Publication Date: August 5, 2000

I hereby grant to the Educational Resources Information Center (ERIC) nonexclusive permission to reproduce and disseminate this document as indicated above. Reproduction from the ERIC microfiche or electronic media by persons other than ERIC employees and its system contractors requires permission from the copyright holder. Exception is made for non-profit reproduction by libraries and other service agencies to satisfy information needs of educators in response to discrete inquiries.
### III. DOCUMENT AVAILABILITY INFORMATION (FROM NON-ERIC SOURCE):

If permission to reproduce is not granted to ERIC, or, if you wish ERIC to cite the availability of the document from another source, please provide the following information regarding the availability of the document. (ERIC will not announce a document unless it is publicly available, and a dependable source can be specified. Contributors should also be aware that ERIC selection criteria are significantly more stringent for documents that cannot be made available through EDRS.)

<table>
<thead>
<tr>
<th>Publisher/Distributor:</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Address:</td>
<td></td>
</tr>
<tr>
<td>Price:</td>
<td></td>
</tr>
</tbody>
</table>

### IV. REFERRAL OF ERIC TO COPYRIGHT/REPRODUCTION RIGHTS HOLDER:

If the right to grant this reproduction release is held by someone other than the addressee, please provide the appropriate name and address:

<table>
<thead>
<tr>
<th>Name:</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Address:</td>
<td></td>
</tr>
</tbody>
</table>

### V. WHERE TO SEND THIS FORM:

Send this form to the following ERIC Clearinghouse:

University of North Carolina at Greensboro  
ERIC/CASS  
201 Ferguson Building  
PO Box 26171  
Greensboro, NC 27402-6171

However, if solicited by the ERIC Facility, or if making an unsolicited contribution to ERIC, return this form (and the document being contributed) to:

ERIC Processing and Reference Facility  
4483-A Forbes Boulevard  
Lanham, Maryland 20706

Telephone: 301-552-4200  
Toll Free: 800-799-3742  
FAX: 301-552-4700  
e-mail: ericfac@inet.ed.gov  
WWW: http://ericfac.piccard.csc.com

EFF-088 (Rev. 2/2000)