Evaluating Psychological Interventions: Efficacy, Effectiveness, Client Progress, and Cost

Robert J. Flynn
University of Ottawa

Before commenting on Hiebert's (1997 [this issue]) paper, I think it is useful to summarize his argument, as accurately and succinctly as possible, in the following propositional form (1) Fiscal restraint has resulted in an increased emphasis within counselling, on accountability, results, and evaluation. (2) To cope successfully with the challenge of accountability, counsellors will have to make evaluation an integral part of the counselling process. (3) Doing so will require an expanded definition of acceptable evidence that includes informal evidence, and encourages counsellors to use informal measures to document client change, such as checklists, the "life line" technique, portfolios, observation forms, cognitive mapping, self-monitored data, authentic assessment, and performance assessment. (4) This new approach to evaluating counselling will need to be collaborative and proactive, involving agreement among all stake holders—funders, special interest groups, managers, counsellors, and clients—on the nature of the service provided, the approach to evaluation, and the evidence that will indicate success and permit stake holders to recognize the effectiveness and value of the service rendered. (5) To guide evaluation practice, agencies will need a workable and integrated counselling-evaluation model that includes a policy statement that identifies stake holders (including funders, special interest groups, clients, significant others, counsellors, supervisors, managers, and coordinators); defines evaluation roles, responsibilities, time frames, and acceptable evidence of success; delineates intervention factors; describes agency factors; and defines communication factors. (6) Finally, counselling needs to emphasize both process, the traditional focus of counselling training and practice, and outcome, a hitherto neglected aspect that needs to be evaluated through built-in procedures.

Critique of Hiebert's Argument

In the present context, I prefer the term *psychological interventions* to that of *counselling*. The former is more general and, as used here, considers counselling and psychotherapy as virtually identical processes that pose the same evaluative questions.

Several of Hiebert's assertions strike me as correct, and I wish to acknowledge the overall utility of his important contribution. I am in essential agreement with propositions 1, 2, and 6 (as formulated above).

Furthermore, if *researchers* were added to Hiebert's lists of stakeholders in propositions 4 and 5, I would also readily agree with them as well. I disagree, however, with Hiebert's third proposition, which is central to his argument. I shall thus concentrate on it here. In doing so, I shall propose what I believe is a more comprehensive evaluation model and a better and more feasible alternative for assessing client progress.

Hiebert's third proposition seems problematic to me, for at least three reasons. First, I think it exaggerates the importance of informal evaluation methods and measures, compared with formal ones. Second, the suggested need to include evaluation as an integral part of counselling (with which I strongly agree) in no way requires, in my opinion, a primary reliance on informal procedures and measures to document client change. On the contrary, formal (i.e., standardized and psychometrically sound) *and* feasible methods and instruments already exist that practitioners are beginning to use to evaluate the status of their clients before, during, and after psychological interventions. Informal methods and instruments may sometimes be useful *adjuncts* to such formal procedures but, in my opinion, they cannot be credible substitutes for them. (For the sake of clarity, let me also state my view that any method or instrument that possesses adequate standardization, norms, reliability, and validity should be considered to be formal rather than informal. Such tools may also be feasible for routine use in clinical decision-making, but they are not automatically so.)

My third criticism of Hiebert's proposition 3 is that his underlying evaluation model seems incomplete. While agreeing with him that the monitoring of individual client progress is crucial, this aspect is but one of the four that I see as essential in a comprehensive approach to evaluation. As the title of my paper suggests, a comprehensive model has to assess not only client progress but also the efficacy, effectiveness, and cost of psychological interventions.

Finally, concerning Hiebert's argument as a whole, I believe that he does not place enough emphasis on the links that need to exist between researchers and practitioners. It is surprising, for example, that his listing of stake holders (in propositions 4 and 5) does not include *researchers*. Convincing assessments of the worth of psychological interventions, however, will require long-term strategic partnerships between practitioners and researchers (Goldfried & Wolfe, 1996). Using mainly formal, well standardized methods, members of these two groups will need to collaborate in the gathering of reliable and valid data on the efficacy, effectiveness, client progress, and costs associated with a wide range of psychological interventions. Examples of such collaboration are beginning to appear, in places such as Ontario and Pennsylvania (Goldfried & Wolfe, 1996).

A Comprehensive Four-Dimensional Approach to Evaluating Psychological Interventions

A recent special issue of the *American Psychologist* (October, 1996) was devoted to the outcome assessment of psychotherapy. Most of the articles in it are directly relevant to the questions raised in Hiebert's paper (which, it should be noted, he had already submitted to the *Canadian Journal of Counselling* before the special *American Psychologist* issue appeared). In preparing my paper, I have relied heavily on the research and reflections contained in this "state of the art" issue of the *American Psychologist*, which synthesizes many of the advances in the outcome evaluation of psychological interventions appearing in the recent literature.

There are four basic questions that may be asked about any psychological intervention and that need to be distinguished carefully from one another (see Howard, Mora, Brill, Martinovich, & Lutz, 1996, Newman & Tejeda, 1996): (1) an *efficacy* question—does the intervention work under the special experimental, controlled conditions of the psychological laboratory? (2) an *effectiveness* question—does it work in practice, under real-world conditions? (3) a *client progress* question—is it working for this particular client? And (4) a *cost* question—how expensive is it? Hiebert's model focuses almost exclusively on the third (client-progress) question without adequate attention to the other three.

Clinical scientists focus on the first (efficacy) question, asking whether a new intervention produces better results than some commonly used intervention or control condition (Howard et al., 1996). The usual method is the randomized clinical trial (RCT), which is designed to have optimal internal validity so that any mean differences among treatment conditions may be attributed to the treatment conditions rather than to some extraneous cause. RCTs use random assignment of clients to treatment conditions, clear specification of the treatments in the form of detailed treatment manuals, selection of clients according to strict inclusion and exclusion criteria, and monitoring of the integrity with which the interventions are actually delivered. RCTs maximize internal validity but often at the expense of external validity, i.e., the degree to which findings may be generalized to other clients, practitioners, or settings (Howard et al., 1996; Seligman, 1996a).

Mental health service researchers concentrate on the second (effectiveness) and fourth (cost) questions, asking whether the new intervention produces good outcomes in the real-world settings of clinics, service agencies, and private-practice offices (Howard et al., 1996), and at what cost (Knapp, 1995; Yates, 1996). Although service researchers may use RCTs, their preferred method is often the naturalistic quasi-experiment in which assignment to comparison groups (e.g., clients with good vs.

poor outcomes) is not random. A good example is the recent *Consumer Reports* retrospective survey (Consumer Reports, 1995; Seligman, 1995, 1996a, 1996b), which concluded that psychotherapy worked very well in the real world. Effectiveness studies often have strong external validity (i.e., generalizability to other clinicians, settings, and clients) but weak internal validity (i.e., observed outcomes may be due to pre-treatment differences on variables other than the independent variable of interest). Replication of the results from such effectiveness studies is thus essential.

Practising clinicians are most interested in the third (client-progress) question of how well an intervention is working for a particular client. The focus is on what is happening to the individual client during (and not only after) the intervention. This, as Hiebert rightly insists, is the most important concern of the practitioner. Fortunately, recent client-oriented research has made possible the systematic assessment of client progress by means of formal, well standardized, reliable, and valid methods.

Monitoring Client Progress through Individual Client Profiling

I do not share Hiebert's belief that practitioners will be able to generate credible evidence of client progress and the worth of counselling through the use of mainly informal evaluation methods. In my opinion, a better and more feasible alternative is to be found in the client-profiling method developed by Howard and his colleagues (Howard et al., 1996). This technique grew out of a dosage model of psychotherapeutic effectiveness, according to which a lawful relationship exists between the log of the number of intervention sessions and the normalized probability of client improvement (Howard, Kopta, Krause, & Orlinsky, 1986). This log-normal model, according to which more and more sessions are needed to produce additional client gains, suggested that psychotherapeutic improvement occurs in three sequential and causally related phases (Howard, Lueger, Maling, & Martinovich, 1993). These are *remoralization* of the client's sense of well-being (typically a rapid phase that takes only a few sessions), *remediation* of symptoms (a longer phase that refocuses the client's coping skills on the obtaining of symptomatic relief), and *rehabilitation* of life functioning (the most gradual phase and one that helps the client learn new ways of handling problem-causing relationship patterns, work habits, or personal attitudes; Howard et. al., 1996).

According to this three-phase model, different interventions and outcomes are appropriate during different phases of counselling, with the outcome criteria for the successive phases being subjective well-being, symptoms, and life functioning, respectively. Standardized scales for measuring each outcome have been developed (for descriptions, see Howard, Brill, Lueger, O'Mahoney, & Grissom, 1995; Howard, Orlinsky,

& Lueger, 1995; and Sperry, Brill, Howard, & Grissom, 1996). An instrument called the Mental Health Index (MHI) serves as an overall measure of outcome and consists of the sum of the subjective well-being score, the current symptom total score, and the current life functioning total score. The MHI has good internal consistency (.87) and test-retest reliability (.82), discriminates well between distressed and non-distressed individuals, and has norms based on over 6,500 clients (Howard et al., 1996).

Plotting the course of an intervention with a particular client on the MHI is straightforward, with periodic assessments being made before, during, and after the intervention. The system includes an individualized criterion of success against which each client's progress can be assessed, namely, the individual client's *expected* progress, given his or her initial clinical characteristics (e.g., the severity and chronicity of the presenting problem, and the client's confidence that the intervention will help). The client-profiling system allows an *expected* MHI score to be generated for each intervention session, which can then be compared graphically with the client's *observed* MHI session scores.

As this new type of client-focused evaluative research increases and is linked with growing knowledge derived from research on the other three dimensions (efficacy, effectiveness, and cost), we may expect additional standardized client-monitoring systems to emerge (see Newman & Tejeda, 1996). The routine use of client profiling in clinical settings would bring several benefits. First, a practitioner could compare a client's actual progress with his or her expected progress to evaluate the appropriateness of the intervention, to plan further intervention, or to request a clinical consultation when the client was not progressing as rapidly as expected (Howard et al., 1996). Second, service agencies would have a feasible means of allocating clients to more or less experienced clinicians, based on the different expected rates of progress of clients who differ in their presenting characteristics. Third, different interventions, clients, practitioners, and agencies could be compared, based on clients' progress during interventions and on their final outcomes. Fourth, collaboration between practitioners and researchers would be encouraged: the former would have a better appreciation of the practical utility of research, and the latter would have a keener understanding of the need to focus on client-progress issues and to relate them to research on the other key dimensions of efficacy, effectiveness, and cost.

In conclusion, use of the formal method of client profiling, with its emphasis on monitoring client status before, during, and after intervention, together with improved data from efficacy, effectiveness, and cost studies, could finally begin to provide the convincing evidence of the worth of counselling that Hiebert and others desire (e.g., Lambert & Cattini-Thompson, 1996). The use of such procedures would also be responsive to the recurring calls for standardized outcome assessment

that are encountered in the counselling literature (Sexton, 1996, provides but the latest example) and would avoid the serious problems related to reliability, validity, and, ultimately, credibility that I believe are inherent in informal methods and measures. The latter should play at best a secondary, adjunctive role in the evaluation of psychological interventions.

References

Consumer Reports. (1995, November). Mental health: Does therapy work? 734-39.

Goldfried, M. R., & Wolfe, B. E. Psychotherapy practice and research: Repairing a strained alliance. *American Psychologist, 51,* 1007-16.

Hiebert, B. (1997). Integrating evaluation into counselling practice: Accountability and evaluation intertwined. *Canadian Journal of Counselling, 31*(2), 99-113.

Howard, K. I., Brill, P. L., Lueger, R. J., O'Mahoney, M. T., & Grissom, G. R. (1995). *Integra outpatient tracking assessment.* Philadelphia: Compass Information Services, Inc.

Howard, K. I., Kopta, S. M., Krause, M. S., & Orlinsky, D. E. (1986). The dose-effect relationship in psychotherapy. *American Psychologist, 41,* 159-64.

Howard, K. I., Lueger, R. J., Maling, M. S., & Martinovich, Z. (1993). A phase model of psychotherapy: Causal mediation of outcome. *Journal of Consulting and Clinical Psychology, 61,* 678-85.

Howard, K. I., Mora, K., Brill, P. L., Martinovich, Z., & Lutz, W. (1996). Evaluation of psychotherapy: Efficacy, effectiveness, and patient progress. *American Psychologist, 51,* 1059-64.

Howard, K. I., Orlinsky, D. E., & Lueger, R. J. (1995). The design of clinically relevant outcome research: Some considerations and an example. In M. Aveline & D. A. Shapiro (Eds.), *Research foundations for psychotherapy practice* (pp. 3-47). Sussex, England: Wiley.

Knapp, M. (Ed.). (1995). *The economic evaluation of mental health care.* London: Arena, Aldershot.

Lambert, M. J., & Cattini-Thompson, K. (1996). Current findings regarding the effectiveness of counselling: Implications for practice. *Journal of Counseling and Development, 74,* 601-08.

Newman, F. L., & Tejeda, M. J. (1996). The need for research that is designed to support decisions in the delivery of mental health services. *American Psychologist, 51,* 1040-49.

Seligman, M. E. P. (1995). The effectiveness of psychotherapy: The *Consumer Reports* study. *American Psychologist, 50,* 965-74.

Seligman, M. E. P. (1996a). Science as an ally of practice. *American Psychologist, 51,* 1072-79.

Seligman, M. E. P. (1996b). A creditable beginning. *American Psychologist, 51,* 1086-88.

Sexton, T. L. (1996). The relevance of counseling outcome research: Current trends and practical implications. *Journal of Counseling and Development, 74,* 590-600.

Sperry, L., Brill, P. L., Howard, K. I., & Grissom, G. R. (1996). *Treatment outcomes in psychotherapy and psychiatric interventions.* New York: Brunner/Mazel.

Yates, B. T. (1996). *Analyzing costs, procedures, process, and outcomes in human services.* Thousand Oaks, CA: Sage.

About the Author

Robert J. Flynn is Director of the Clinical Psychology Program at the University of Ottawa.

Address correspondence to: Robert J. Flynn, Centre for Psychological Services, School of Psychology, University of Ottawa, 11 Marie Curie, Vanier Hall (6th floor), Ottawa, ON K1N 6N5. Electronic mail may be sent via Internet to rflynn@uottawa.ca.