Commentary on <u>Pragmatic Case Studies and Evidence-Based Treatment: Research and Clinical Applications of a Computerized Outcomes Management System</u>

Technology to Support the Clinical Management of Psychotherapy Cases: Commentary on the Polaris-MH

ROBERT J. LUEGER [a,b]

[a] Department of Psychology, Marquette University, Milwaukee, WI

[b] Correspondence concerning this article should be addressed to Robert J. Lueger, Associate Dean for Academic Affairs, College of Arts and Sciences, Department of Psychology, Marquette University, Milwaukee, WI 53201-1881.
Email: robert.lueger@marquette.edu

_____

ABSTRACT

Advances in the technology of measurement and informatics have potential for supporting the clinical management of individual cases of psychotherapy. The Polaris-MH system described by Grissom and Lyons (2006) is one example of these advances. The potential utility of this and similar systems is discussed in light of research on feedback to therapists and of prediction of treatment outcomes.

Key words: psychotherapy outcomes; mental health treatment; learning systems; feedback in psychotherapy

_____

Behavioral scientists, particularly psychologists, have a long history of developing measures of constructs useful for assessing changes over time that might be attributable to interventions. The development of a useful measure involves identification of a target construct; the selection of thoughts, feelings, and behaviors that sample the construct; iterative refinement of the item content of the measure to ensure homogeneity of the construct and heterogeneity of the associated behaviors, thoughts, and feelings; evidence of convergence with measures of the same or similar constructs; and either predictive or criterion validity. Collection of data from representative samples enables normative comparisons, and knowledge of the measure's sensitivity to change informs judgments about the meaningfulness of differences in scores collected at successive time points separated by elements of the intervention. The availability of electronic technology for immediate scoring and feedback permits use of knowledge gleaned from such a measure to be used to correct a treatment that is in progress. Ideally this facilitates a "smart system" in which corrections to the treatment result in a pattern of improvement that more closely approximates an ideal, or at least acceptable, pattern given the clinical characteristics of the patient, the availability of appropriate treatment techniques, and therapist

expertise. The Polaris-MH system described by Grissom and Lyons (2006), is an example of a measure designed to work within a "smart system."

The late Kenneth Howard (Howard & Lueger, 1991; Howard, Brill, Lueger, & O'Mahoney, 1992) was among the first to contribute to the systematic measurement of patient progress in psychotherapy. Using empirical data from aggregated cases of individual psychotherapy, Howard and colleagues developed an omnibus measure of mental health status that reflected three domains of patient improvement — recovery of hope, amelioration of symptoms, and rehabilitation of deficits in social and role functioning. Psychotherapy researchers, recognizing the potential of monitoring change during psychotherapy, convened the Vanderbilt Conference (Strupp, Horowitz, & Lambert, 1994) to revisit a recurrent idea, namely, the possibility of identifying and reaching consensus on the contents of a core battery that could be used universally in studies of psychotherapy process and outcome. But consensus was elusive, and psychotherapy researchers began to develop their own measures to track progress in psychotherapy. Undoubtedly the most successful of these was the 45-item Outcome Questionnaire (OQ-45) developed by Michael Lambert (Lambert, Hansen, Umpress, Lunnen, Okiishi, Burlingame, & Reisinger, 1996). In the United Kingdom, the CORE measure (Evans, Mellor-Clark, Margison, Barkham, Audin, Connell, & McGrath, 2000) was developed on the Howard tri-dimensional model to assess treatment outcomes in the English health care system. Meanwhile, Howard and his colleagues continued the development of their system, albeit under different labels (INTEGRA, COMPASS, TeaM), and aggregated thousands of cases of patients/clients treated with psychotherapy or counseling. The Polaris-MH system represents a second-generation development within the Howard tradition.

THEORETICAL OR CRITERION-INFORMED CONTENT

Two distinct approaches have emerged in the construction of measures and systems for monitoring progress in psychotherapy. As detailed in the report by Grissom and Lyons (2006), the Polaris-MH system was constructed to reflect *a priori* theoretical constructs within the omnibus measure, Behavioral Health Status. This theoretical approach follows the work of Howard, but also is represented in the CORE that is used in the United Kingdom. By contrast, the OQ-45 is more of a criterion-defined measure. Lambert reviewed clinical records of improved and unimproved psychotherapy cases to identify aspects of symptom improvement and social and role functioning that were noted to have improved in treatment. Although there are similarities in content to the three dimensions of the Howard model, the OQ-45 typically is reported as a global score of mental health functioning. Partly because the Polaris-MH measure was designed to reflect three distinct dimensions of potential change, the internal consistency of the global score ($r = .83$) is less than that of the single construct OQ-45 ($r = .94$). This has implications for the measure's sensitivity to change.

How do the theoretically- and criterion-informed approaches to measurement compare? Grissom and Lyons (2006) observed a correlation of 0.87 between the Behavioral Health Status score of the Polaris and the total score of the OQ-45, which is a high concurrent validity

correlation. Obviously these two measures of overall mental health status capture elements of a similar construct. What clinical value might the tri-dimensional approach of the Polaris add to an assessment of patient response in psychotherapy? Both measures appear to assess symptom remediation with equal efficiency and success, and should yield quite similar results when compared in clinical use. The theoretically-informed dimensions of the Polaris-MH system — remoralization, remediation of symptoms, and rehabilitation of functioning — have the potential for specific feedback to therapists provided that improvement in these dimensions can be related to overall treatment change. According to the Phase Model, remoralization (improvement in Subjective Well-Being on the Polaris) is a precondition for symptom remediation. Therefore, failure to improve as expected on the symptom dimension might be queried regarding any change on Subjective Well-Being, and lacking evidence of such improvement, feedback to the therapist could be directed at helping the patient feel more hopeful. Given evidence that symptoms have improved, the Functional Disability dimension of the Polaris might capture further improvements, especially in longer-term psychotherapy, that the global score of the OQ-45 does not. Although the Phase Model (Howard, Lueger, Maling, & Martinovich, 1993) is a theoretically rich approach to mental health treatment outcomes, its demonstrated utility in feedback schemes is more promising than actual at this point in the development of "smart systems."

RELIABILITY AND SENSITIVITY TO CHANGE

Reliability of measurement is critically important for an index that purportedly captures changes attributable to treatments. Grissom and Lyons (2006) convincingly argue the case for using an estimate of internal consistency rather than the index of stability over time that is used with trait measures. Lambert has adopted the same approach with the OQ-45. The range of reliabilities for the Polaris-MH measure is deemed "acceptable" by the authors, and certainly is representative of other construct measures. However, with a coefficient alpha of $r = .83$, a patient's score must increase one standard deviation (10 T-score points) on the omnibus measure, Behavioral Health Status (BHS), to be considered "statistically reliably improved" at a 95% confidence level. If one recalls the evidence of meta-analyses of psychotherapy outcomes (e.g., Wampold, 2001), the typical or average outcome of a person treated with psychotherapy is about one standard deviation better than an untreated person across a wide range of measures. Thus the BHS measure would seem to be better suited to assess pre-post changes than in-process changes. By contrast, the Symptoms dimension of the Polaris-MH is sufficiently sensitive to detect changes of a magnitude of one-half standard deviation (5 T-score points). Thus it would seem to be the measure of choice in monitoring change during treatment.

Sensitivity to change over time is another relevant issue for monitoring instruments such as the Polaris-MH. Although reliability and sensitivity are psychometrically intertwined, what is at issue is an instrument's capacity to detect changes that truly happen. For example, an internally consistent state measure is more likely to detect change over time than is an internally consistent trait measure. A monitoring instrument that includes items reflecting traits is less

likely to be sensitive to changes than one that contains more items measuring states. One way to assess the sensitivity of items in the overall scale is to examine the capacity of each item to distinguish between the known improved and non-improved status of psychotherapy patients. For example, an unpublished analysis of the 68 items of the "COMPASS" measure (the predecessor of the BHS measure) that I conducted revealed that changes on fewer than 40 items accounted for most of the variance distinguishing improvers and non-improvers.

FEEDBACK FROM MONITORING SYSTEMS

Peterson's (1991) model of disciplined inquiry is a useful scheme for examining the potential contributions of a feedback system such as the Polaris-MH. In the Peterson model, client characteristics (A), the guiding conception (rubric) of treatment (B), and therapist experience (C) contribute through an assessment (D) of the client's mental health status to produce a case formulation and treatment plan (E). These in turn lead to therapeutic action (F). Monitoring (G) of patient progress while in treatment can lead to a reformulation of treatment (E) if treatment progress is unsatisfactory, or to affirmation of the client (A) if progress is satisfactory.

Michael Lambert and his group (Lambert, Hansen, & Finch, 2001) have contributed most to the research on the impact of feedback by focusing on potential treatment failures. Lambert's color-coding scheme is both descriptive and predictive (a yellow signal cautions that the patient is not progressing, whereas a red signal is predictive that the patient is on course to be worse off after psychotherapy than before treatment began). Lambert works on the principle that 10-15% of psychotherapy patients will statistically decline in overall mental health status as a result of the psychotherapeutic intervention. In randomized control studies, Lambert (cited in Grissom & Lyons, 2006) has shown that therapist knowledge of patient failure (a red signal) leads to the provision of more treatment and subsequently fewer treatment failures. The research data on feedback about treatment success has been mixed: in some studies, therapist knowledge of success (a green signal) has led to earlier terminations, but in other studies this has not been replicated.

The second research group steadily contributing to knowledge of the impact of feedback on treatment is that led by Hans Kordy (Percevic & Kordy, 2003). Using a probabilistic, stochastic "Random Walk" model, this group has shown that future improvement in the course of psychotherapy is best predicted by knowledge of the patient's current status relative to normative comparisons of improved patients or "normal" non-patients. Simply said, if the patient has room for improvement, given enough therapy that improvement is likely to be observed.

By contrast, the Howard group seeks to identify the projected pattern of change given the clinical characteristics of the client or patient. Current progress is assessed against expected treatment progress given these clinical characteristics. Apart from the statistics involved in calculating such projected trajectories of change, this process is arguably more similar to the intuitive approach of practicing therapists who assess success as a function of this therapy with

this patient given these clinical characteristics. There currently are not studies from Grissom and Lyons and others working in the Howard model to evaluate the impact of this kind of feedback on treatment outcomes.

PREDICTING OUTCOMES OF PSYCHOTHERAPY

How valid are the predictions of outcome based on observed change scores of monitoring instruments such as the Polaris-MH scales? An assumption of clinical informatics or "smart systems" is that the information conveyed is reliably and validly predictive of a desired outcome. The Polaris-MH measures follow the Expected Treatment Response approach of the Howard group (cf. Lueger, Howard, Martinovich, Lutz, Anderson, & Grissom, 2001). It is important to note that empirical research has found that the error margins of projected paths progressively increase from the point of prediction. Thus a prediction of a client's termination status that is made at the initial session of psychotherapy is likely to have broader error bands than one made after a half-dozen sessions of psychotherapy (Lutz, Rafaeli-Mor, Howard, & Martinovich, 2002).

The new field of clinical informatics in psychotherapy requires evidence of predictive validity to make predictions about treatment outcome that have practical usefulness. Purely *descriptive* feedback statements on progress (e.g., "The patient's score is not reliably different from his initial score at intake," or "The patient is following the expected pattern of response given her clinical characteristics.") do not venture a prediction of treatment outcome, and thus they leave predictive interpretation in the mind of the therapist. *Predictive* feedback statements (e.g., "The patient has a high probability of treatment failure," or, "Given the patient's response in therapy to date, you can expect that the patient will be a treatment success.") are actuarially based and require evidence of validity to lend confidence in their prediction. There is much to do to identify the base rates and the sensitivity and specificity of predictors of outcome. Predictive validity is established for classes or groups of similar individuals. The degree to which any particular individual of that group will deviate from the expected response of the group is unknown at the time of prediction. Post-treatment analyses by the therapist and researcher might lead to a refinement of prediction parameters to further increase predictive validity in subsequent cases, but there still is that element of unknown deviation from the group central tendency that might render the prediction for a specific individual more or less accurate. In other words, as therapists we attempt to predict outcome for a single individual, for whom there is no error term (except by replication of the same situation and that is a rare event).

To illustrate these points, consider some empirical findings involving the Expected Treatment Response. Lutz and colleagues (Lutz, Martinovich, & Howard 1999) have shown that a patient who falls below the 25th percentile (-.675 s.d.) of all patients with similar clinical characteristics is on a path for treatment failure. If the patient falls in that zone twice in the first ten sessions, she/he has a 90% chance of failure (high specificity). Unfortunately, this sign is observed in only 10% of the patients who eventually fail (low sensitivity) in psychotherapy (Lueger, 2003). So, for a group of 100 similar patients expected to respond with a failure pattern,

only 10 will show a sign that is highly specific for failure; moreover, one of those ten will not fail, and we don't know with certainty, for an individual patient, whether this is one of the ten. Only afterwards, can we do a post-mortem analysis to see if the predictive algorithm can be improved with additional information from the aberrant psychotherapy participant.

CONCLUSION

The Polaris-MH system merges measurement, the technology of immediate scoring and feedback, and information from large data sets to inform the therapist, and potentially the patient, about the ongoing success of the psychotherapeutic treatment. There are many issues -- involving psychometrics, predictive validity, and feedback utility -- that importantly need more development in this system. We look forward with interest and excitement to such development of this system into a state-of-the-art, clinical informatics approach.

REFERENCES

Evans, C., Mellor-Clark, J., Margison, F., Barkham, M., Audin, K., Connell, J., & McGrath, G. (2000). CORE: Clinical outcomes in routine evaluation. *Journal of Mental Health, 9*, 247-255.

Grissom, G. & Lyons, J. (2006). Pragmatic case studies and evidence-based treatment: Research and clinical applications of a computerized outcomes management system. *Pragmatic Case Studies in Psychotherapy [Online]*, Vol. 2(3), Article 1. Available: http://hdl.rutgers.edu/1782.1/pcsp_journal.

Howard, K.I., & Lueger, R.J. (1991, July). *Monitoring change in psychotherapy.* Paper presented at the annual meeting of the Society for Psychotherapy Research, Lyon, France.

Howard, K.I., Brill, P., Lueger, R.J., & O'Mahoney, M. (1992). *The Integra Outpatient Tracking Assessment*. Radnor, PA: Integra, Inc.

Howard, K.I., Lueger, R.J., Maling, M.S., & Martinovich, A. (1993). A phase model of psychotherapy: Causal mediation. *Journal of Consulting and Clinical Psychology, 61*, 678-685.

Lambert, M., Hansen, N., & Finch, A. (2001). Patient-focused research: Using treatment outcome data to enhance treatment effects. *Journal of Consulting and Clinical Psychology, 69*, 159-172.

Lambert, M., Hansen, N., Umpress, V., Lunnen, K., Okiishi, J., Burlingame, G.M. & Reisinger, C.W. (1996). *Administration and scoring manual for the OQ-45.2 (Outcome Questionnaire).* Stevenson, MD: American Professional Credentialing Services, LLC.

Lueger, R.J. (2003, June). Predicting treatment outcomes and gains from feedback on symptom changes. Paper presented at the international meeting of the Society for Psychotherapy Research, Weimar, Germany.

Lueger, R.J., Howard, K.I., Martinovich, Z., Lutz, W., Anderson, E.A., & Grissom, G. (2001). Assessing treatment progress of individual patients with expected treatment response models. *Journal of Clinical and Consulting Psychology, 69*, 150-158.

Lutz, W., Martinovich, Z., & Howard, K.I. (1999). Patient profiling: An application of random coefficient regression models to depicting the response of a patient to outpatient psychotherapy. *Journal of Consulting and Clinical Psychology, 67*, 571-577.

Lutz, W., Rafaeli-Mor, E., Howard, K. I., & Martinovich, Z. (2002). Adaptive modeling of progress in outpatient psychotherapy. *Psychotherapy Research, 12*, 427-443.

Percevic, R., & Kordy, H. (2003). The Random Walk Model: Predicting outcomes of psychotherapy. Paper presented at the international meeting of the Society for Psychotherapy Research, Rome, Italy.

Peterson, D. (1991). Connection and disconnection of research and practice in the education of professional psychologists. *American Psychologist, 40*, 441-451.

Strupp, H.H., L. M. Horowitz, & M. J. Lambert, (1994.). *Measuring patient changes in mood, anxiety, and personality disorders: Toward a core battery.* Washington, D.C.: American Psychological Association.

Wampold, B.E. (2001). *The great psychotherapy debate.* Mahwah, N.J.: Lawrence Erlbaum Associates.