

The Psychotherapy Dose Effect in Naturalistic Settings Revisited: Response to Gray

Gray's response (this issue) to Hansen, Lambert, & Forman (2002) suggests that, due to a number of limitations that exist in the original paper, the effectiveness of psychotherapy in naturalistic settings may be underestimated. This response to Gray addresses three of the limitations raised, including: (1) using clinical significance, it is difficult to adequately capture treatment gains made by people who are already within the functional range on an outcome instrument when they initiate treatment; (2) the end point data used to assess treatment outcome in the Hansen et al. (2002) paper does not fall at the end of the last session of treatment; and (3) for people beginning treatment in the most severely dysfunctional range on an outcome instrument, the effect size for treatment is quite large. In this response we demonstrate that, even if these limitations are accounted for, the amounts of treatment delivered and the subsequent response to treatment clearly lags what is observed in clinical trials research.

Gray (this issue) raises some good points—some that seem warranted and others that do not. Three specific issues are dealt with here. First, using clinical significance criteria does pose a problem when dealing with patients who initiate treatment but are already in the functional range (Tingey, Lambert, Burlingame, & Hansen, 1996). This is a key reason that achieving Reliable Change should be considered meaningful improvement, even though this departs from the traditional “gold standard” of recovery as the outcome of interest in CS methodology. Further, patients initiating treatment already in the functional range on traditional outcome measures show less change on average than more distressed patients (Hansen & Lambert, 2003). Depending on the clinical setting, the percentage of patients who initiate treatment in this range can be quite variable—in Hansen et al. (2002), the proportion of those in the functional range at the beginning of treatment ranged from 18.3% to 42.4%, with an overall average of 36.2% across sites. If we were to assume that these patients are not appropriate to consider in an outcome analysis and

therefore drop them from the data set, then our rates of improvement and recovery do increase—for recovery from 14.1% to 21.7%, and for improvement from 20.9% to 21.6%. So while this brings the total treatment response rate for naturalistic settings to 43.3%, it does not double the rate as was suggested by the letter author, and this figure is still quite a bit lower than the 67.2% improved/recovered observed in RCTs. It should also be remembered that 19.8% of patients who began treatment in the functional range did improve. In fact, a survival analysis of similar data (Hansen & Lambert) reveals that when only looking at those who begin treatment in the functional range, 50% of those who receive at least 10 (or 11, depending on site) sessions of therapy reliably improve. When considering that the average length of treatment reported by Hansen et al. was four sessions or less, it is clear that most patients who receive treatment in clinical settings do not have adequate exposure to treatment to achieve this level of change. It should also be pointed out that low dosage is not a recent phenomenon that is solely the consequence of managed care; it has been true for decades (Garfield, 1994). Nevertheless managed care organizations have a responsibility to try to overcome utilization problems.

Second, Gray correctly points out that the majority of the end-point assessment data used by Hansen et al. (2002) occurs at the beginning of the last session of therapy, and therefore the effect of this last session (and occasionally, the last few sessions) is not included in the computation of outcome status. This could have an impact on the rates of improvement and recovery, particularly in treatments that are as brief as those observed by Hansen et al. Referring again to Hansen and Lambert (2003), however, a survival analysis (that estimates the consequences of receiving additional sessions) of similar data does not suggest that one or two additional sessions of therapy brings median treatment lengths (from four up to six sessions) into the range that would produce adequate treatment response (15 to 19 sessions for 50% of patients to recover, 8 to 12 sessions for 50% of patients to improve).

Finally, the RCI in clinical significance methodology can be thought of as the amount of change needed to achieve a significant effect—corrected for measurement error. For the OQ-45, an effect size of about 0.64 is needed to achieve reliable change. Considering the table

presented by Gray, only patients who begin treatment in the severely distressed range have an average effect size that reaches this level. Patient change scores have to cross a cutoff point between dysfunction and functional distributions to be considered recovered, and if every patient in the sample had a change score equal to 0.86 standard deviations, and this sample has a distribution similar to general psychiatric patients, it is likely that less than half would cross the cutoff point (assuming that the short form of the OQ performs similarly to the full). According to this table then, even in the severely distressed range, where on average every patient changes enough to reliably improve, the rate of recovery may still lag behind what is observed in RCTs. Further, it is likely that those with higher distress scores receive more sessions of treatment. It is not just where one starts treatment that affects outcome, but also how much treatment one receives.

In summary, the estimates presented by Hansen et al. (2002) are rough at best, and may indeed (as suggested by Gray) underestimate typical treatment response to some degree. Even if this is the case, however, the extremely large national data that were analyzed suggest that outcomes seen in naturalistic settings are relatively low and don't correspond to outcomes observed in clinical trials research. Gray correctly points out that any comparison between RCTs and naturalistic clinical settings is problematic, and that given the less controlled nature of "real world" clinical work, where patient problems and corresponding clinical approaches are likely to be more varied, it would stand to reason that outcomes would trail those of RCTs, which probably provide the upper bounds of effective practice. It also appears that effectiveness studies that transfer treatment protocols from RCTs to nonresearch settings with less carefully selected patients can replicate the level of outcomes found in RCTs, but only when treatment doses are replicated (Merrill, Tolbert, & Wade, 2003; Wade, Treat, & Stuart, 1998). Patient, practitioner, and third-party payer expectations about treatment duration and expected outcome need realigning, and need to be

based on a more firm empirical foundation. Further research on this topic is needed.

REFERENCES

- Garfield, S. L. (1994). Research on client variables in psychotherapy. In A. E. Bergin & S. L. Garfield (Eds.), *Handbook of Psychotherapy and Behavior Change* (4th ed., pp. 190–228). New York: Wiley.
- Hansen, N. B., Lambert, M. J., & Forman, E. M. (2002). The psychotherapy dose–response effect and its implications for treatment delivery services. *Clinical Psychology: Science and Practice, 9*, 329–343.
- Hansen, N. B., & Lambert, M. J. (2003). An evaluation of the dose–response relationship in naturalistic treatment settings using survival analysis. *Mental Health Services Research, 5*, 1–12.
- Merrill, K. A., Tolbert, V. E., & Wade, W. A. (2003). Effectiveness of cognitive therapy for depression in a community mental health center: A benchmarking study. *Journal of Consulting & Clinical Psychology, 71*, 404–409.
- Tingey, R. C., Lambert, M. J., Burlingame, G. M., & Hansen, N. B. (1996). Assessing clinical significance: Proposed extensions to method. *Psychotherapy Research, 6*, 109–123.
- Wade, W. A., Treat, T. A., & Stuart, G. L. (1998). Transporting an empirically supported treatment for panic disorder to a service clinic setting: A benchmarking strategy. *Journal of Consulting and Clinical Psychology, 66*, 231–239.

Nathan B. Hansen
Department of Psychiatry,
Yale University School of
Medicine
Michael J. Lambert
Department of Psychology,
Brigham Young
University
Evan M. Forman
Department of Psychology,
Drexel University

DOI: 10.1093/clipsy/bpg051