Psychology and Psychotherapy: Theory, Research and Practice (2011), 84, 389–405

The British Psychological Society

www.wileyonlinelibrary.com

Piloting a practice research network: A 12-month evaluation of the Human Givens approach in primary care at a general medical practice

William Andrews[1]*, Elspeth Twigg[2], Takuya Minami[3] and Gina Johnson[4]

[1]Nottingham Trent University, UK
[2]CORE Information Management Systems, UK
[3]Northeastern University, Massachusetts, USA
[4]Stopsley Group Practice, UK

Objective. To investigate the effectiveness of the Human Givens (HG) approach to the management of emotional distress in a primary care setting.

To investigate whether or not the use of a shorter version (i.e., CORE-10) of a well-established psychometric instrument (i.e., Clinical Outcome in Routine Evaluation (CORE) CORE-outcome measure, CORE-OM) for sessional data collection is feasible for large-scale implementation of a practice research network (PRN).

Design. All clients who chose to opt into assessment for treatment with three accredited HG therapists following referral for management of psychological distress, primarily anxiety and depression, by General Medical Practitioners (GPs) or GP practice nurses working in a primary care general medical practice over a 12-month period were included.

Methods. The primary outcome measures were the CORE-OM and CORE-10. Pre–post effect sizes (Cohen's *d*) were calculated using pre, post, and pooled standard deviations to facilitate comparison with previously published studies. Mixed-design analysis of variance (ANOVA) was used to look at differences in pre- and post-treatment symptoms and potential treatment effects based on type of termination and gender. Observed intent-to-treat pre–post effect size using the CORE-OM was also benchmarked against data from Clark *et al.* (2009) improving access to psychological therapies (IAPT) pilot site data. Results obtained using CORE-OM were compared with those obtained using CORE-10 to evaluate the feasibility of using the CORE-10 for routine use in real-world clinical settings.

Results. Pre- to post-treatment changes measured with the CORE-OM and CORE-10 suggested that the therapy was highly effective, with clients remaining in treatment to completion demonstrating the greatest benefit. Reliable change and recovery rates

*Correspondence should be addressed to William Andrews, 77 Osborne Road, Sheffield S11 9BA, UK (e-mail: wandrews22@mac.com).

DOI:10.1111/j.2044-8341.2010.02004.x

comparisons between the CORE-OM and CORE-10 indicated that the CORE-10 is a viable alternative to the CORE-OM. Result of the benchmarking indicated that the observed pre–post effect size was clinically equivalent to IAPT data published by Clark *et al.* (2009).

Conclusions. Although replications are warranted as the current investigation is a pilot study, the HG approach appears to be an effective treatment. CORE-10 is a satisfactory generic sessional assessment to use in place of the 34-item CORE-OM. Use of a shorter yet reliable outcome measure is likely to increase assessment completion rates. PRNs appear to be a suitable mechanism to establish treatment effectiveness across a wide range of treatments in different settings.

Practice research networks (PRNs) are defined as collaborations of practicing providers who commit to using their work settings as laboratories for practice-based knowledge generation (McMillen, Lenze, Hawley, & Osborne, 2009). Although PRNs have been well established in primary medical care since the 1950s, they have only been introduced into mental health since the mid-1990s, with the first example being the American Psychiatric Institute for Research and Education's Practice Research Network (www.psych.org/MainMenu/Research/APIRE.aspx) formed in 1993. Others such as the American Psychological Association [APA] (http://apapracticenet.net) and the American Association of Marriage and Family Therapists [AAMFT] (www.aamft.org) have followed suit. While these provide examples of guild-specific PRNs where much of their work has involved member surveys, the Pennsylvania Practice Research Network (Borkovec, Echemendia, Ragusea, & Ruiz, 2001) was established with the emphasis on investigation of practice, with a particular focus on client progress during outpatient therapy. In the UK, the Art Therapists' PRN (Huet & Evans, 2008) was established to provide a means for clinicians to be involved in practitioner-led projects without undergoing specialist research training (Parry, Castonguay, Borkovec, & Wolf, 2010). Barkham and colleagues (Barkham *et al.*, 2001; Mullen, Barkham, Mothersole, Bewick, & Kinder, 2006) have long promoted practice-based research and evaluation for the exploration of the richness of clinical data obtained across practice settings using identical measures. This work has led to the development of national benchmarks with regard to service delivery in primary care, measuring treatment acceptance rates, recovery and improvement rates, waiting times, and other metrics (Bewick, Trusler, Mullen, Grant, & Mothersole, 2006; Mullen, Barkham, Mothersole, Bewick, & Kinder, 2006). The value of providing service providers with feedback on service user progress session-by-session through clinical feedback systems has also been advocated (Duncan, Miller, & Sparks, 2004; Lambert, Harmon, Slade, Whipple, & Hawkins, 2005). In sum, PRNs provide a rich environment to establish effectiveness of treatments delivered in the real world.

Human Givens

The Human Givens (HG) approach (Griffin & Tyrrell, 2004, 2008) can best be understood as an integrative bio/psycho/social approach to the promotion of mental health where empirically grounded clinical interventions are employed to assist service users to get their physical and emotional needs met. These needs should be met in healthy balanced ways through the use of individuals' own resources, both received as a part of their genetic makeup and further developed through learning, development and life experience. The approach is fundamentally grounded in the organizing idea that human beings have both needs and resources and, like all living entities that thrive, require these

needs to be met and these resources to work effectively. With respect to the delivery of mental health service, this central backbone of the HG approach acts as a compass for both service provider and service user in directing the therapeutic endeavour. These organizing ideas are also based on more recent evidence regarding the effects of high emotional arousal (LeDoux, 1998; Phan, Wager, Taylor, & Liberzon, 2002), how the brain processes trauma (Brewin & Holmes, 2003), how mindfulness affects the activity in the brain (Kabat-Zinn, 2003; Seigel, 2007), and HG theory on the role of REM sleep and dreaming (Griffin & Tyrrell, 2006). Fosha, Siegel, and Solomon (2009) reference the most current research on the interconnectedness of the mind, brain, body, emotions, and social relationships by respected authorities from a wide range of scientific disciplines. While the organizing ideas may have different emphasis and move away from prescriptive diagnosis, the HG approach still uses many empirically grounded clinical interventions (Salkovskis, 2002).

The HG approach incorporates many of the principles of Cognitive Behavioural Therapy (CBT), using brief, time-limited interventions, focussing on symptom relief and, primarily, on the current situation rather than the client's past, with the exception of trauma, where a form of trauma-focussed imaginal exposure is typically used (Muss, 1991). However, the HG approach has a different focus from that of CBT – on identification of unmet emotional needs and resolution of unhelpful established emotional patterns, rather than on questioning and testing unhelpful thoughts and beliefs – and uses a broader range of therapeutic techniques, recognizing that there is no one method that will cope effectively with the diversity of human cognition, emotions, and experiences.

Training in the HG approach is on a peripatetic basis, usually spread over some years, with attendance on an initial programme of discrete training days followed by an intensive residential programme. Final assessment and accreditation is by way of video submission and case report on therapy work. Therapists make their own arrangements for supervision according to recommended guidelines of the HG Institute (www.hgi.org.uk).

Following a systematic literature search, Corp, Tsaroucha, and Kingston (2008) suggested that most evidence related to the HG approach is expert opinion or anecdotal (Griffin & Tyrrell, 2004, 2008). Other phase I pre–post studies have been published (e.g., Guy & Guy, 2009) but these have not been in peer-reviewed journals. While the treatment delivered and the techniques employed for the commonly presenting problems in primary care is broadly in line with guidelines outlined by the National Institute for Health and Clinical Excellence (NICE, www.nice.org.uk), until the commencement of the current pilot project, there had been no systematic attempt at evaluation of the approach.

Awareness of the Department of Health (DH) recommendations (NIMH(E), 2004) with respect to outcome measurement led to our decision that the most suitable manner to set about ensuring high quality data collection on the employment of the HG approach was to establish a PRN (HGIPRN- www.hgiprn.org), where practitioners would commit to gathering contextual and outcome data, using the same instruments, with every service user at every session. We also believed that it was fundamentally important that employment of any measures within the therapeutic context must make sound clinical sense and be of benefit to both service provider and service consumer, as well as facilitating clinical supervision. We chose to use a proprietary Internet-based software system, CORE Net, developed by CORE IMS (www.coreims-online.co.uk) so practitioners widely dispersed across the UK could all enter data that could then be

collated centrally and reported on. We decided to pilot the PRN over a 12-month period at a general medical practice in Luton. The practice has around 7,000 patients and is located in a suburban area of low deprivation. Three accredited HG therapists were employed 1 day a week each to accept all referrals from the doctors (GPs) and nurse practitioners (NPs) that previously had been made to an outside agency for Primary Care Counselling. The project was approved by the Evidence Based Practice Committee of Luton Teaching Primary Care Trust.

Method

Participants

Demographics

The mean age of the 124 clients was 42.8 years ($SD = 12.7$, $n = 123$), with a median of 44 years (range 17–81 years). Thirty-six clients were male (29.0%) and 88 were (71%) female. One hundred and seven (86.3%) of the clients were described as White British; one (.8%) as Black/Black British African; one (.8%) as White Irish; one (.8%) as White (English/European) and two (1.6%) as Asian/Asian British Indian. No data on ethnicity were available for 12 clients (9.7%). At their first session, 66 (53.2%) of clients were on medication. There was no significant difference between the proportions of males and females who were taking medication (males 61%, females 50%; $z = .927$, $p > .05$).

Measures

Clinical outcome measures were collected at every contact between service provider and service user, with the outcome measures offered near the beginning of sessions. To ensure clinical utility, rather than use disorder-specific measures, we chose to use the CORE-outcome measure (CORE-OM; Evans *et al.*, 2000; Barkham *et al.*, 2001) at pre-treatment, post-treatment (where possible, even when the ending was unplanned), and follow up. The CORE-OM is a 34-item questionnaire addressing domains of subjective well-being; symptoms (anxiety, depression, physical problems, trauma); functioning (general functioning, close relationships, social relationships) and risk (risk to self, risk to others). Items are scored on a 0–4 Likert-type scale rated over the past week and the clinical score is the mean of all items multiplied by 10. Forms with up to three items missing are considered valid. The recommended cutoff between clinical and non-clinical populations is 10 (Connell & Barkham, 2007). The internal consistency of the CORE-OM has been reported as $\alpha = .94$ and the 1-week test–retest reliability as Spearman's $\rho = 0.90$ (Evans *et al.*, 2002). On a sessional basis, we used CORE-10 (Connell *et al.*, 2007), a 10-item version of the CORE-OM that uses items drawn from the CORE-OM in order to minimize the load on the client. CORE-10 has been demonstrated to be sensitive to change and to correlate very highly with CORE-34 with $r = .94$ in a clinical sample. Reported internal consistency was also high ($\alpha = .82$; Connell *et al.*, 2007). Both the CORE-OM and CORE-10 include a mixture of positively and negatively framed, and high- and low-intensity items. Internal consistency for the current sample was $\alpha = .93$ (CORE-OM; $n = 101$) and $\alpha = .80$ (CORE-10; $n = 114$) using their initial assessments, which was equivalent to the high reliability reported in previous studies.

CORE therapy assessment (TA) and end of therapy (EoT) forms are designed to be completed by practitioners and give information on presenting problems, demographics, etc. Practitioners were asked to complete a CORE TA form for every referral assessed and a CORE EoT form if the client attended for two or more sessions.

Table 1. Severity and duration of depression and anxiety symptoms at initial assessment

	Severity				
Duration of symptoms	Mild	Moderate	Mod-severe	Severe	Total
Depression					
<6 months	0	9	5	4	18
6–12 months	2	1	8	4	15
>12 months	1	1	3	7	12
Recurring/continuous ('enduring')	0	10	15	22	47
Total	3	21	31	37	92
Anxiety					
<6 months	0	3	7	3	13
6–12 months	0	2	9	1	12
>12 months	1	1	6	7	15
Recurring/continuous ('enduring')	0	3	26	30	59
	1	9	48	41	99

Diagnoses, severity, and duration of presenting problems

The CORE system uses a cutoff of 10 on the CORE-OM or 11 on the CORE-10 to indicate whether a client falls within a clinical population. Of the 121 clients (out of 124 total clients) who had a valid CORE-OM initial assessment, 107 (88.4%) clients were above the clinical cutoff on the CORE-OM at pre-treatment; using the CORE-10 (embedded in the CORE-OM), 110 (91.7%) of the 120 clients who had a valid initial assessment were above the clinical cutoff. Based on the CORE TA and EoT forms, 92 (74.2%) clients were experiencing some degree of depression before treatment and 99 (79.8%) were experiencing some symptoms of anxiety. Table 1 below gives a detailed breakdown of the reported severity and duration of anxiety and depression. Of note is the fact that by far the largest group for both problems is those who are experiencing symptoms on a recurring/continuous basis (i.e., 'enduring' symptoms; depression: $n = 47$ of 92 [51.1%] clients; anxiety: $n = 59$ of 99 [59.6%] clients). Figure 1 provides the participant flowchart for the current study.

Interventions

Mindful of the extensive research into general change mechanisms (Bergin & Garfield, 1994) as well as focussing on all the common factors widely recognized for their importance, such as establishing a therapeutic relationship and validating the service user's experience, therapists working with the HG approach place great emphasis on resource activation (Gassmann & Grawe, 2006). As described above, the therapeutic focus tends to be on the identification of the unmet needs of the service user and on the development of strategies using solution-focussed problem solving, goal setting, mental rehearsal, and skills training to help the service user work towards being better able to get these needs met. Interventions include but are not limited to the identification of positive experiences and personal attributes, goal setting, establishing a collaborative treatment plan, reframing, challenging thoughts, activity scheduling, use of guided imagery, behavioural activation, controlled breathing, progressive muscle relaxation, setting session agendas, graded exposure (imaginal and *in vivo*), development of interpersonal relationships, motivational interviewing, and training in the practice of

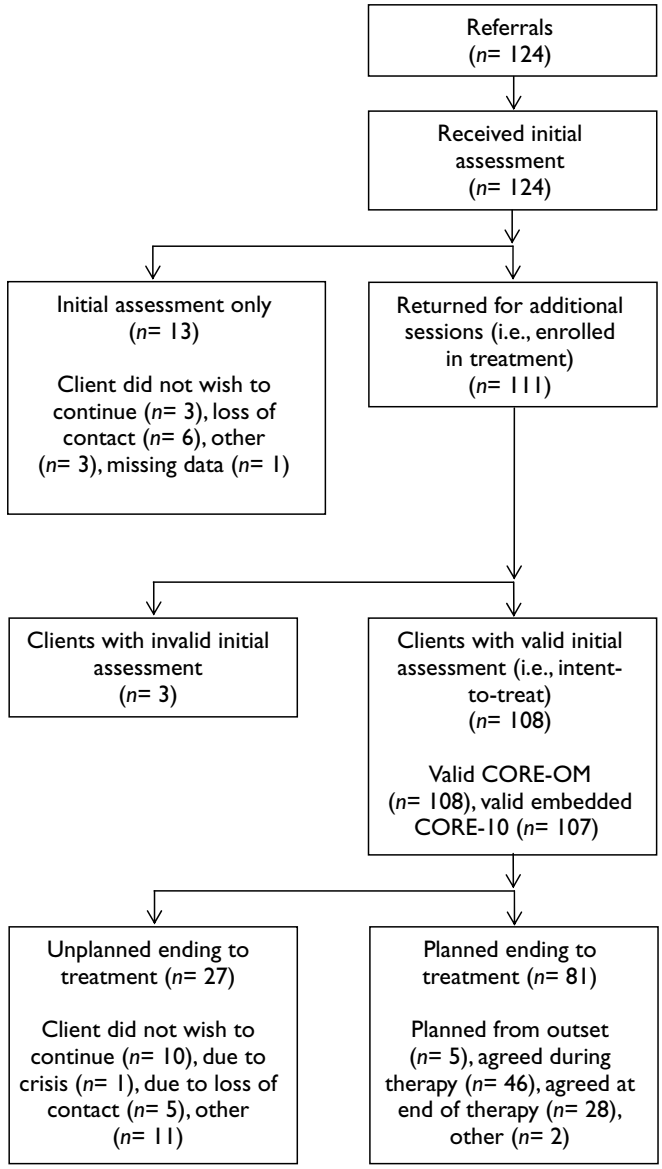

Figure 1. Participant enrolment and treatment flowchart.

mindfulness. Aware of the uniqueness of each individual, the integrative nature of the approach allows for a great deal of flexibility on the part of the therapist.

Procedures

The GPs and practice nurses were introduced to the HG approach at a practice meeting 1 month prior to commencement of the service. They referred service users to the service administrator when they felt that the service user would likely benefit from the service. The administrator explained in more detail about the service and provided some literature on the HG approach. Service users were invited to think it over and phone in

when they felt ready to make an appointment if they felt it was right for them. When the service user phoned in, he/she was assigned to whichever of the three practitioners had space available at a suitable time. Therefore, although not randomized to the three practitioners, no client or therapist preferences were taken into consideration when assigning clients to practitioners. The three therapists involved in this study had monthly personal supervision and quarterly group supervision. We gathered process and outcome data on the 12-month cohort of service users in Luton referred by the GPs and treated between April 2007 and March 2008. The clinical protocol was designed so that the data collection was compulsory for every service user. The administrator oversaw data gathering to ensure 100% compliance and entered all the data into CORE Net. A total of 124 clients were referred and assessed by the three practitioners.

Analyses

Because of the high correlation between the CORE-OM and the CORE-10 ($r = .94$), it was deemed appropriate to use the CORE-10 scores (either embedded in CORE-OM or not) to measure change in clients who entered but did not complete therapy and where it proved impossible to obtain a final CORE-OM. The gains in terms of pre–post data collection were judged to outweigh the data losses incurred as a result of using the shorter measure, in addition to being able to calculate a pre–post effect size for the intent-to-treat sample. This decision was also influenced by our intent to assess whether or not the implementation of the CORE-10, which is a shorter version of the CORE-OM, is practically feasible as a measure to monitor treatment outcomes in the real world. Therefore, the pre–post effect size using the CORE-10 was deemed the main outcome, although parallel analyses were also conducted using the CORE-OM as confirmation. Pre–post effect sizes were calculated using pre, post, and pooled standard deviations to facilitate comparisons with previously published studies depending on which standard deviations the comparison studies utilized.

Within-study effectiveness

As the initial assessment, whether or not changes in clinical symptoms occurred between pre- and post-treatment was tested. A mixed-model repeated-measures ANOVA was conducted to assess, simultaneously, the changes in clinical symptoms between pre- and post-treatment (within-subjects), potential effect of gender on treatment effect (between-subjects), and potential effect of planned *versus* unplanned ending on treatment effect (between-subjects).

Benchmarking

To assess whether or not the pre–post treatment effect size obtained in the current study meets acceptable standards of treatment effectiveness in the real world, the current data were tested using the IAPT data published in Clark *et al.* (2009) as the benchmark (Minami, Serlin, Wampold, Kircher, & Brown, 2008; Minami, Wampold, Serlin, Hamilton, Brown, & Kircher, 2008; Minami, Wampold, Serlin, Kircher, & Brown, 2007). Specifically, Clark *et al.*'s pre–post data as assessed by the CORE-OM ($n = 232$; Doncaster and Newham sites combined) were used as the benchmark so as to keep the outcome measures identical between the benchmark and the current data. Pre–post effect size was calculated using the formula recommended by Becker (1988) using the pre-treatment standard deviation and correcting for bias in estimation as established by Hedges (1981).

The effect size was then statistically tested against Clark *et al.*'s pre–post data as the benchmark with a one-tailed range-null hypothesis testing (Serlin & Lapsley, 1985, 1993) using a clinical equivalence margin of 10% of the benchmark following Minami, Wampold *et al.* (2008). In other words, if the pre–post treatment effect size estimated from the current data is within 10% of the IAPT data by Clark *et al.*, it was considered that the effect size was clinically equivalent to that of the benchmark.

Reliable change and recovery

Following Jacobson and Truax (1991; also Lambert & Ogles, 2009), change scores of clients who began at the clinical range and had valid post-treatment scores ($n = 97$ using CORE-10; $n = 79$ using CORE-OM) were categorized based on (a) whether or not their change was reliable (i.e., 'reliable change') and (b) whether or not clients moved from clinical range to non-clinical range (i.e., 'recovered'). Reliable change was measured using previously defined criteria of a change in CORE-OM of 5 or more points (CORE Systems Group, 1998; Jacobson & Truax, 1991) or a change in CORE-10 score of 6 or more points (Connell & Barkham, 2007). In addition, reliable change and recovery rates obtained using the CORE-10 and the CORE-OM was compared with one another to evaluate the feasibility of utilizing the CORE-10 for routine assessment in the real world instead of the longer CORE-OM.

Results

Data completeness

Of the 124 clients with at least one session data, 111 (89.5%) clients returned for at least one additional session (i.e., considered 'entered treatment'). There was no evidence of differences among those who did and did not enter treatment based on gender or initial symptom severity (overall, depression, or anxiety; $p = .073$–$.633$).

Of those who entered treatment, 108 clients (97.2%) had valid pre-treatment CORE-OM data and 90 clients (81.2%) had valid post-treatment CORE-OM data. One hundred and seven clients (96.4%) had pre-treatment CORE-10 data (as embedded in CORE-OM) but one client had missing items within the 10 items. Therefore, valid post-treatment data were obtained from 106 clients (99.1%).

Clinical outcomes

Within-study effectiveness

Table 2 shows pre- and post-treatment scores for all clients who entered therapy, split according to the type of therapy ending (i.e., planned or unplanned). A mixed-model repeated-measures ANOVA was conducted to test the change in clinical symptoms from pre- to post-treatment (within-subjects) as well as potential effects of gender and planned *versus* unplanned ending on treatment effectiveness (between-subjects). Pre–post outcome using the CORE-10 was considered the primary indicator due to the larger sample size and the high correlation with the CORE-OM.

The within-subjects effect assessing change between pre- and post-treatment showed that participants significantly improved between the two timepoints (CORE-10: $F[1,102] = 75.75$, $p < .001$; CORE-OM: $F[1,86] = 76.33$, $p < .001$). No evidence of treatment effect based on gender were indicated (CORE-10: $F[1,102] = 2.10$, $p = .151$; CORE-OM: $F[1,86] = .470$, $p = .495$). However, there were evidence of differential treatment effect based on whether or not termination was planned or not (CORE-10:

Table 2. Pre- and post-treatment scores and effect sizes (Cohen's *d*) on CORE-10/CORE-OM by type of therapy ending and benchmark CORE-OM data from Clark *et al.* (2009)

		Pre-treatment		Post-treatment		*d*		
	Ending	*n*	*M* (*SD*)	*n*	*M* (*SD*)	Pre-*SD*	Pooled-*SD*	Post-*SD*
CORE-10	Planned	80	18.53 (7.23)	80	8.79 (6.61)	1.33	1.39	1.45
	Unplanned	27	22.04 (6.24)	26	14.54 (9.05)	1.13	0.91	0.77
	Total	107	19.41 (7.13)	106	10.19 (7.65)	1.27	1.23	1.18
CORE-OM	Planned	81	16.78 (6.33)	76	7.73 (5.58)	1.43	1.51	1.60
	Unplanned	27	20.17 (5.16)	14	11.55 (8.62)	1.28	1.01	0.83
	Total	108	17.63 (6.21)	90	8.33 (6.25)	1.41	1.40	1.40
Clark *et al.* (2009)	Total	232	18.5 (6.0)	232	11.1 (7.3)	1.22	1.10	1.00

Note. Effect sizes (Cohen's *d*) are calculated using data available from clients who had both valid pre- and post-treatment CORE-10 or CORE-OM assessment. Data from Clark *et al.* (2009) combine Doncaster and Newham sites. Their data were not separated into planned and unplanned termination.

$F[1,102] = 11.75$, $p = .001$; CORE-OM: $F[1,86] = 4.48$, $p = .037$). As expected, clients whose terminations were planned had significantly better clinical outcomes than those who had unplanned terminations. Figure 2 shows the mean pre- and post-treatment CORE-10 scores for the current study.

Benchmarking

Using the CORE-OM as the outcome measure and the pre-treatment *SD* as the deviation to calculate the pre–post treatment effect size, the IAPT data from Clark *et al.* (2009) was $d = 1.22$ ($n = 232$; Doncaster and Newham sites combined; Table 2). As their study included all clients who had valid scores at both pre- and post-treatment, the same criterion was applied with the pre–post treatment effect size in the current data. The 90 clients who had valid CORE-OM assessments at both pre- and post-treatment were included regardless of whether or not their termination was planned. The observed

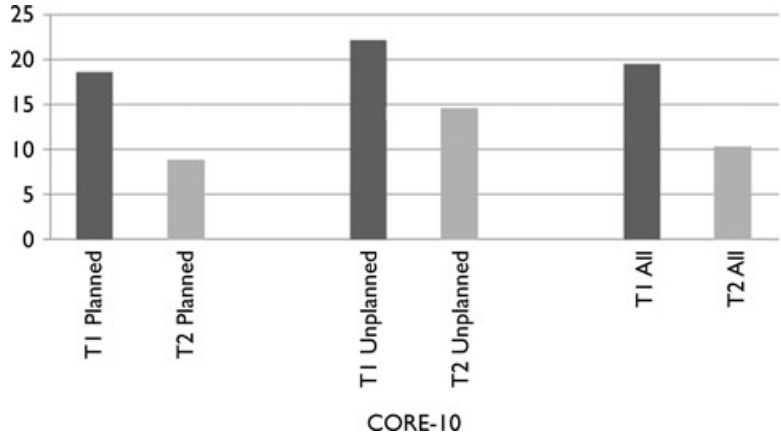

Figure 2. Luton mean pre- and post-CORE-10 scores (HGIPRN – Luton study) planned ($n = 80$), unplanned ($n = 26$), all ($n = 106$) (pre-treatment = T1, post-treatment = T2).

Table 3. Reliable change and recovery rates based on CORE-OM/CORE-10

Measure	Reliable change	Recovery		Total
		No	Yes	
CORE-10 ($n = 97$)	No	22 (22.7%)	3 (3.1%)	25 (25.8%)
	Yes	19 (19.6%)	53 (54.6%)	72 (74.2%)
	Total	41 (42.3%)	56 (57.8%)	97 (100.0%)
CORE-10 (of $n = 79$ below with valid CORE-OM)	No	16 (20.3%)	1 (2.9%)	17 (21.5%)
	Yes	13 (16.5%)	49 (60.9%)	62 (78.5%)
	Total	29 (36.7%)	50 (63.3%)	79 (100.0%)
CORE-OM ($n = 79$)	No	16 (20.3%)	4 (5.1%)	20 (25.3%)
	Yes	11 (13.9%)	48 (60.8%)	59 (74.7%)
	Total	27 (34.2%)	52 (65.8%)	79 (100.0%)

pre–post treatment effect size, similarly using the pre-treatment *SD*, thus resulted in $d = 1.41$ ($n = 90$; Table 2). The one-tailed non-central t test using a 10% margin of clinical equivalence was statistically significant ($t = 13.37$, $ncp = 10.40$, $df = 89$, $p = .019$). Therefore, not only was the observed pre–post effect size of $d = 1.41$ larger than that of IAPT data by Clark *et al.* ($d = 1.22$), the statistically significant result against the benchmark shows promise of clinical equivalence with the benchmark for future large-scale implementations of the HGIPRN.

Reliable change and recovery

Clients' change scores were categorized based on reliable change and recovery (Table 3). Indices based on the CORE-10 showed that 54.6% of the participants with pre-therapy scores in the clinical range demonstrated reliable change and also recovered. As the participants with valid CORE-10 data include those that had unplanned terminations, the percentage is an estimate of an intent-to-treat population.

CORE-OM post-treatment data were not available for clients with unplanned terminations of therapy. However, 60.8% of participants who completed HG therapy demonstrated reliable change and recovery. Comparing this percentage against Mullen *et al.* (2006), the rate of clients who had reliable change and also recovered fell well within the top quartile (75th percentile in Mullin *et al.* was 58% [CI = 55.4 to 65.9]).

Comparing the indices obtained with the CORE-10 and the CORE-OM with the same 79 participants (Table 3), utilization of CORE-10 appeared to be very similar as expected from the very high correlation between the two ($r = .94$). Therefore, it appears that large-scale implementation of routine assessment is defensible using the shorter version (i.e., 10 items) of the scale rather than the full 34-item version, which would most likely aid in increased data collection as strongly emphasized by Clark *et al.* (2009).

Outcomes by duration of presenting problem (anxiety and depression) pre-therapy

In order to investigate differences in scoring and moving to recovery based on presenting problems, we used the practitioner ratings of severity and duration of symptoms of depression or anxiety/stress to identify clients with enduring problems (greater than 6 months of moderately severe/severe symptoms). Table 4 shows clients' scores based on severity of symptoms at assessment.

Table 4. Pre–post effect sizes based on 'enduring' depression and anxiety status

		Pre-treatment		Post-treatment		*d*		
	'Enduring'	*n*	*M* (*SD*)	*n*	*M* (*SD*)	Pre-*SD*	Pooled-*SD*	Post-*SD*
Depression								
CORE-10	Yes	50	20.99 (6.94)	49	12.12 (8.74)	1.24	1.09	0.98
	No	57	18.04 (7.07)	57	8.54 (6.19)	1.33	1.41	1.51
CORE-OM	Yes	50	19.59 (6.09)	42	9.96 (6.76)	1.48	1.40	1.32
	No	58	15.93 (5.85)	48	6.90 (5.44)	1.42	1.48	1.54
Anxiety								
CORE-10	Yes	69	20.27 (7.17)	68	11.48 (8.07)	1.20	1.13	1.06
	No	38	17.85 (6.87)	38	7.89 (6.31)	1.42	1.48	1.55
CORE-OM	Yes	70	18.54 (6.30)	58	9.00 (6.44)	1.39	1.38	1.36
	No	38	15.94 (5.76)	32	7.10 (5.77)	1.45	1.45	1.45

Note. Effect sizes (Cohen's *d*) are calculated using data available from clients who had both valid pre- and post-treatment CORE-10 or CORE-OM assessment using pre-treatment, pooled, and post-treatment standard deviations.

Discussion

The current pilot study assessed the feasibility of a large-scale implementation of the Human Givens Practice Research Network (HGIPRN) by conducting a within-study analysis of pre–post treatment effect, benchmarking, and reliable change and recovery rates. General results from the Luton pilot indicated that there were significant changes in clients' symptoms using both the CORE-10 and CORE-OM from pre- to post-treatment. Observed pre–post treatment effect size was clinically equivalent to benchmark data (i.e., Clark *et al.*, 2009) and also attained reliable change and recovery rates that were comparable to results observed in a previous study (i.e., Mullen *et al.*, 2006). In all, these results appear promising for expanding the HGIPRN.

Our decision to ask clients to complete a CORE-10 at all but their first and last session meant that we were able to obtain data for any service users attending two or more sessions, regardless of whether their treatment ending was planned or unplanned. From the standpoint of establishing a PRN, it is absolutely crucial that data are collected consistently. The very high correlation between the CORE-OM and CORE-10 scores, as well as the near-identical categorization of reliable change and recovery rates, indicate that it would be reasonable to use CORE-10 items embedded in the CORE-OM as their initial assessment but to use the CORE-10 on a sessional basis. Although use of the CORE-OM on a routine basis would be preferable to use of CORE-10 from a psychometric standpoint, results show that routine use of CORE-10 on a sessional basis could be a better choice given the high rate of compliance using the CORE-10 and the high correlation between the two measures.

There is controversy about the importance of missing post-treatment data (Clark *et al.*, 2009; Clark, Fairburn, & Wessely, 2008; Richards & Suckling, 2009; Stiles, Barkham, Mellor-Clark, & Connell, 2008a,b). Session-by-session monitoring, with almost complete data collection, proved to be an important achievement in the IAPT demonstration sites of Doncaster and Newham (Clark *et al.*, 2009; Richards & Suckling, 2009). Targets set by the DH for IAPT nationally are 90% second-timepoint data and 50% 'moving to recovery' (IAPT, 2009). In the current study, 100% of service users completed a pre-treatment assessment and 100% of service users who entered treatment (attended more

than one session) filled in at least one subsequent assessment, although, as in any data collection, in one case the measure proved invalid because not enough questions were answered.

There does not appear to be consensus among researchers regarding the best methodology to adopt for calculation of pre–post treatment effect size where repeated measures are used. For example, Minami, Wampold *et al.* (2008), following Becker (1988), chose to use the *SD* of the pre-treatment score on the basis that it is less influenced by repeated measurement and/or treatment. In contrast, Clark *et al.* (2009) used the pooled *SD*; Richards and Suckling (2009) used the post-treatment *SD* 'as it provides the most conservative estimate of effect size' (p. 382). However, the current study shows that the post-treatment *SD* is not always smaller than the pre-*SD*, and therefore, using the post-treatment *SD* does not necessarily provide the most conservative estimate. Therefore, our recommendation is to use the pre-*SD*, which, following Becker's logic, is unaffected by the repeated measurements or the treatment and thus describes the estimated variability among clients regarding their clinical symptoms prior to treatment in the population. We nevertheless chose to calculate the pre–post effect sizes using all three *SD*s because the published IAPT effect sizes have been calculated using the pooled *SD* (Clark *et al.*, 2009) and the post-treatment *SD* (Richards & Suckling, 2009).

Our effect sizes for completers were in a range from $d = 1.33$–1.56, depending on whether one used CORE-OM or CORE-10 scores and which *SD* was used. On the other hand, the effect size for unplanned endings dropped to a range between $d = 0.77$ and 1.23. Given that premature dropout was associated on average with much poorer outcomes, it stresses the importance of collecting data on the type of therapy ending and further investigations into the cause and prevention of premature termination. However, it is interesting to note that the larger variability in outcomes for clients with premature termination (i.e., larger post *SD*) suggests that some clients may have dropped out because they already received what they needed, whereas others may have dropped out because therapy may not have been a good fit for them or other life events arose. Further research is needed to explore these issues in much more depth.

Problem duration and medication

Natural recovery without treatment is recognized as a strong possibility where psychological distress is of recent onset (Clark *et al.*, 2009). However, 60% of the Luton cohort reported moderately severe or severe depressive symptoms of duration in excess of 6 months, with 69% reporting moderately severe or severe symptoms of anxiety of more than 6 months' duration. In fact, only 15% of the cohort reported problems of less than 6 months' duration. Again, this compares very favourably with the cohorts reported in the pilot IAPT sites by Clark *et al.* Studies of depression and anxiety where recruited cases have a duration of 6 months or greater tend to report very low natural recovery rates, in a range of 5–20% (Clark *et al.*, 1994, 1998; Posternak & Miller, 2001). While 53% of the cohort was on medication during their treatment, there was no difference in outcomes between those who were on medication and those who were not. It seems reasonable to conclude that, in the majority of cases in the Luton study, the treatment offered was strongly associated with service user progress towards recovery. However, as there have been reports of medication effects over and above therapy (e.g., $d = 0.15$ reported in Minami, Wampold *et al.*, 2008), it is important to consider the possibility that for those 53% of the clients, the pre–post treatment effect size may have included medication effects.

Other limitations

Being an observational prospective cohort study with no wait-list control, this study suffers from the same limitations as the IAPT published data (Clark *et al.*, 2009, pp. 9–10) that we utilized as the treatment effect size benchmark. To draw satisfactory causal inference with regard to the effectiveness of the HG approach in primary care, we would need to conduct an RCT study with wait-list control or established treatment arm. While therapists working with the HG approach utilize pragmatic variants of many empirically supported clinical interventions recommended in NICE guidelines, it could be argued that this prospective cohort study cannot be seen as a phase IV study examining the implementation of NICE guidelines, but, rather, as a very robust phase I study examining pre/post therapy effect tested against a benchmark that is accepted by the research community. The study also lacks evidence regarding the correct delivery of treatment and, consequently, evidence of the precise components of the treatment. The profile of the particular service user cohort in terms of social support and ethnic mix would also limit the generalizability of these findings to the primary care population at large.

Implications

When Salkovskis (1995) first described the hourglass model of psychotherapy research, it appears it was not his intention that the randomized clinical trial (RCT) should be seen at the top of an evidence hierarchy. In fact, in 2002, he raised grave concerns about the direction of psychotherapy research: 'the risk inherent in the current practice of evidence based mental health is that the field will degenerate into a parody, a kind of one-dimensional science' (Salkovskis, 2002, p. 4). While Richards and Suckling (2009) do refer to the IAPT Doncaster project as a phase IV prospective cohort study, we note that Clark *et al.* (2009) concluded that 'neither (Doncaster nor Newham) could be described as comprehensive services that implemented the NICE guidelines for the psychological treatment of depression and all the anxiety disorders' (p. 11). This raises the question about the feasibility of actually tightly adhering to NICE guidelines given that naturalistic practice settings are quite different from research environments (Nathan, Stuart, & Dolan, 2000; Wampold, 2001; Westen & Morrison, 2001; Westen, Novotny, & Thompson-Brenner, 2004). Westfall, Mold, and Fagnan (2007) point out that demonstration of efficacy and causality in clinical trials does not automatically suggest effectiveness in the real world of everyday practice. Horn and Gassaway (2007) discuss the emergence in medicine of robust and comprehensive gathering of practice-based evidence for clinical practice improvement as a complementary process for exploring effectiveness of treatments in real-world settings. It may be that the implication is that NICE guidelines should be seen as exactly that – 'guidelines' rather than 'directives' that require a very narrow interpretation – and the NICE evidence hierarchy should broaden its umbrella to include other kinds of robust practice-based evidence of the quality provided by studies such as the current one conducted in real-world settings.

This pilot study acted as a template for the establishment of a PRN as a mechanism to which real-world practice effects could be measured. Obtaining agreement on choice of outcome measures and contextual data to be captured and choosing an Internet-based system that would allow a consistent approach of data entry across practitioners and sites formed the crucial first steps in the process of establishing the PRN. Training was then carried out across the UK at various HG approach peer groups by the first author. Online training was also provided and was made available through a bespoke protected page on

the PRN website (www.hgiprn.org). A robust and rigorous approach was emphasized throughout the training.

It was recognized that gathering of such robust evidence would be greatly facilitated by the more widespread use of practical generic sessional tools such as CORE-10, as well as the more disorder-specific instruments such as GAD-7 in cases and settings that warrant more specific but large-scale assessments. Where outcome measure availability at a second timepoint exceeds 90%, it becomes possible to develop recovery and improvement benchmarks of service and practitioner performance in routine primary care. Once these benchmarks are established in the real world, the applicability of research-based benchmarks established under conditions that are far removed from actual practice becomes questionable (Minami, Wampold *et al.*, 2008).

The HG approach is informed by NICE guidelines and the evidence upon which they are based. This initial pilot study suggests that the approach is a *bona fide* treatment that significantly contributes to assisting service users in primary care to move towards recovery. While accepting the limitations outlined above, the study demonstrates that high quality robust outcome assessment is feasible in a real-world setting and this has indeed set the bar very high for data quality for the HGIPRN as it rolls out across a wide variety of services and service providers. It also provides an example of possibility for other services and *bona fide* treatment approaches as well as creating the possibility of developing new benchmarks for data quality in mapping the service user journey from point of referral to follow-up.

Conclusion

The piloting of the HGIPRN demonstrated that the approach is suitable for training and supporting HG practitioners in gathering robust high quality sessional outcome data for use in evaluation of service user progress in treatment. Early indications suggest that the HG approach is an effective treatment for working with service users presenting with a variety of problems, and particularly anxiety and depression, in primary care settings. Treatment effect sizes compared favourably with those of service users seen in IAPT services. The value of CORE-10 as a more brief sessional measurement tool has been demonstrated. McMillen *et al.* (2009) consider that PRNs have been underused in mental health service research. We believe that this pilot, a mental health treatment effectiveness evaluation utilizing a PRN, encourages the concept of the 'scientist-practitioner' at its fullest and places the service user at its heart.

References

Barkham, M., Margison, F., Leach, C., Lucock, M., Mellor-Clark, J., Evans, C., *et al*. (2001). Service profiling and outcomes benchmarking using the CORE-OM: Toward practice-based evidence in the psychological therapies. Clinical outcomes in routine evaluation-outcome measures. *Journal of Consulting and Clinical Psychology*, *69*, 184–196.

Becker, B. J. (1988). Synthesizing standardized mean-change measures. *British Journal of Mathematical and Statistical Psychology*, *41*, 257–278.

Bergin, A. E., & Garfield, S. L. (Ed.) (1994). *Handbook of psychotherapy & behaviour change*. New York: John Wiley & Sons.

Bewick, B. M., Trusler, K., Mullen, T., Grant, S., & Mothersole, G. (2006). Routine outcome measure completion rates of the CORE OM in primary care psychological therapies and counselling. *Counselling & Psychotherapy Research*, *6*, 50–59.

Borkovec, T. D., Echemendia, R. J., Ragusea, S. A., & Ruiz, M. (2001). The Pennsylvania practise research network and future possibilities for clinically meaningful and scientifically rigorous psychotherapy effectiveness research. *Clinical Psychology: Science and Practice*, *8*, 155–167.

Brewin, C. R., & Holmes, E. A. (2003). Psychological theories of posttraumatic stress disorder. *Clinical Psychology Review*, *23*, 339–376.

Clark, D. A., Fairburn, C. G., & Wessely, S. (2008). Psychological treatment outcomes in routine NHS services: A commentary on Stiles *et al.* (2007). *Psychological Medicine*, *38*, 629–634.

Clark, D. M., Layard, R., Smithies, R., Richards, D. A., Suckling, R., & Wright, B. (2009). Improving access to psychological therapy: Initial evaluation of two UK demonstration sites. *Behaviour Research and Therapy*, *47*(11), 910–920.

Clark, D. M., Salkovskis, P. M., Hackmann, A., Middleton, H., Anastasiades, P., & Gelder, M. (1994). A comparison of cognitive therapy, applied relaxation and imipramine in the treatment of panic disorder. *British Journal of Psychiatry*, *164*, 759–769.

Clark, D. M., Salkovskis, P. M., Hackmann, A., Wells, A., Fennell, M., Ludgate, J., *et al*. (1998). Two psychological treatments for hypochondriasis. A randomised controlled trial. *British Journal of Psychiatry*, *173*, 218–225.

Connell, J., & Barkham, M. (2007). *CORE-10 User Manual, Version 1.1*. CORE System Trust & CORE Information Management Systems Ltd. Rugby, UK.

Connell, J., Barkham, M., Stiles, W. B., Twigg, E., Singleton, N., Evans, O., *et al*. (2007). Distribution of CORE-OM scores in a general population, clinical cut-off points, and comparision with the CIS-R. *British Journal of Psychiatry*, *78*, 127–138.

CORE System Group (1998). *CORE system (information management) handbook*. Leeds, UK.

Corp, N., Tsaroucha, A., & Kingston, P. (2008). Human givens therapy: The evidence base. *Mental Health Review Journal*, *13*, 44–52.

Duncan, B. L., Miller, S. D., & Sparks, J. A. (2004). *The heroic client; a revolutionary way to improve effectiveness through client directed outcome informed therapy*. San Francisco: Jossey-Bass.

Evans, C., Connell, J., Barkham, M., Margison, F., Mellor-Clark, J., McGrath, G., *et al*. (2002). Towards a standardised brief outcome measure: Psychometric properties and utility of the CORE-OM. *British Journal of Psychiatry*, *180*, 51–60.

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

Fosha, D., Siegel, D. J., & Solomon, M. F. (2009). *The healing power of emotion; affective neuroscience, development & clinical practice*. New York: W. W. Norton & Co.

Gassmann, D., & Grawe, K. (2006). General change mechanisms: The relation between problem activation and resource activation in successful and unsuccessful therapeutic interactions. *Clinical Psychology and Psychotherapy*, *13*, 1–11.

Griffin, J., & Tyrrell, I. (2004). *Human givens: A new approach to emotional health and clear thinking*. Chalvington: Human Givens Publishing Ltd.

Griffin, J., & Tyrrell, I. (2006). *Dreaming reality; how dreaming can keep us sane or drive us mad*. Chalvington: Human Givens Publishing Ltd.

Griffin, J., & Tyrrell, I. (2008). *An idea in practice: Using the human givens approach*. Chalvington: Human Givens Publishing Ltd.

Guy, K., & Guy, N. (2009). Psychological trauma. *Counselling at Work* (Spring), 19–22.

Hedges, L. V. (1981). Distribution theory for Glass's effect size and related estimators. *Journal of Educational Statistics*, *6*, 107–128.

Horn, S., & Gassaway, J. (2007). Practice-based evidence study design for comparative effectiveness research. *Medical Care*, *45*, S50–S57.

Huet, V. S. N., & Evans, C. (2008). *Art therapy practice research network*. Retrieved from http://www.baat.org/atprn.html.

IAPT (2009). *Improving Access to Psychological Therapies: Key performance indicators and technical guidance*. Retrieved from http://www.iapt.nhs.uk/2009/03/improving-access-to-psychological-therapies-key-performance-indicators-and-technical-guidance-2008/

Jacobson, N. S., & Truax, P. (1991). Clinical significance: A statistical approach to defining meaningful change in psychotherapy research. *Journal of Consulting and Clinical Psychology*, *59*, 12–19.

Kabat-Zinn, J. (2003). Mindfulness-based interventions in context: Past, present and future. *Clinicasl Psychology, Science and Practice*, *10*, 144–156.

Lambert, M. J., Harmon, C., Slade, K., Whipple, J. L., & Hawkins, E. J. (2005). Providing feedback to psychotherapists on their patients' progress: Clinical results and practice suggestions. *Journal of Clinical Psychology*, *61*, 165–174.

Lambert, M. J. & Ogles, B. M. (2009). Using clinical significance in psychotherapy outcome research: The need for a common procedure and validity data. *Psychotherapy Research*, *19*(4), 493–501.

LeDoux, J. (1998). *The emotional brain*. London: Phoenix; a division of Orion Books Ltd.

McMillen, J. C., Lenze, S. L., Hawley, K. M., & Osborne, V. A. (2009). Revisiting practice-based research networks as a platform for mental health services research. *Administration and Policy in Mental Health*, *36*, 308–321.

Minami, T., Serlin, R. C., Wampold, B. E., Kircher, J. C., & Brown, G. S. (2008). Using clinical trials to benchmark effects produced clinical practice. *Quality & Quantity*, *42*, 513–525.

Minami, T., Wampold, B. E., Serlin, R. C., Hamilton, E. G., Brown, G. S., & Kircher, J. C. (2008). Benchmarking the effectiveness of psychotherapy treatment for adult depression in a managed care environment: A preliminary study. *Journal of Counsulting and Clinical Psychology*, *76*, 116–124.

Minami, T., Wampold, B., Serlin, R. C., Kircher, J. C., & Brown, G. S. (2007). Benchmarks for psychotherapy efficacy in adult major depression. *Journal of Counsulting and Clinical Psychology*, *75*, 232–243.

Mullen, T., Barkham, M., Mothersole, G., Bewick, B. M., & Kinder, A. (2006). Recovery and improvement benchmarks in routine primary care mental health settings. *Counselling & Psychotherapy Research*, *6*, 68–80.

Muss, D. C. (1991). A new techinque for treating post-traumatic stress disorder. *British Journal of Clinical Psychology*, *30*, 91–92.

Nathan, P. E., Stuart, S. P., & Dolan, S. L. (2000). Research on psychotherapy efficacy an effectiveness: Between Scylla and Charybdis? *Psychological Bulletin*, *126*, 964–981.

NIMH(E). (2004). *Organising and delivering psychological therapies: July 2004*. Retrieved from http://www.dh.gov.uk/PublicationsAndStatistics/Publications/PublicationsPolicyAndGuidance/PublicationsPolicyAndGuidanceArticle/fs/en?CONTENT_ID=4086100&chk=72vikq

Parry, G., Castonguay, L., Borkovec, T., & Wolf, A. W. (2010). Practice research networks and psychological services research in the UK and USA. In M. Barkham, G. E. Hardy, & J. Mellor-Clark (Eds.), *Developing and delivering practice-based evidence: A guide for the psychological therapies* (pp. 311–325). Oxford: John Wiley.

Phan, K. L., Wager, T., Taylor, S. F., & Liberzon, I. (2002). Functional neuroanatomy of emotion: A meta-analysis of emotion activation studies in PET and fMRI. *Neuroimage*, *16*, 331–348.

Posternak, M. A., & Miller, I. (2001). Untreated short-term course of major depression: A meta-analysis of outcomes from studies using wait-list control groups. *Journal of Affective Disorders*, *66*, 139–146.

Richards, D. A., & Suckling, R. (2009). Improving access to psychological therapies: Phase IV prospective cohort study. *British Journal of Clinical Psychology*, *48*, 377–396.

Salkovskis, P. M. (1995). Demonstrating specific effects in cognitive and behavioural therapy. In M. Aveline & D. Shapiro (Ed.), *Research foundations for psychotherapy practice*. Chichester: Wiley.

Salkovskis, P. M. (2002). Empirically grounded clinical interventions: Cognitive-behavioural therapy progresses through a multi-dimensional approach to clinical science. *Behavioural and Cognitive Psychotherapy*, *30*, 3–9.

Seigel, D. J. (2007). *The mindful brain*. New York: Norton & Co.

Serlin, R. C., & Lapsley, D. K. (1985). Rationality in psychological research: The good-enough principle. *American Psychologist*, *40*, 73–83.

Serlin, R. C., & Lapsley, D. K. (1993). Rational appraisal of psychological research and the good-enough principle. In G. Keren & C. Lewis (Eds.), *A handbook for data analysis in the behavioral sciences: Methodological issues* (pp. 199–228). Hillsdale, NJ: Lawrence Erlbaum Associates.

Stiles, W. B., Barkham, M., Mellor-Clark, J., & Connell, J. (2008a). Effectiveness of cognitive-behavioural, person-centered, and psychodynamic therapies in UK primary care routine practice: Replication with a larger sample. *Psychological Medicine*, *38*, 677–688.

Stiles, W. B., Barkham, M., Mellor-Clark, J., & Connell, J. (2008b). Routine psychological treatment and the Dodo verdict: A rejoinder to Clark *et al*. *Psychological Medicine*, *38*, 905–910.

Wampold, B. E. (2001). *The great psychotherapy debate: Models, methods, and findings*. Mahwah, NJ: Lawrence Erlbaum Associates.

Westen, D., & Morrison, K. (2001). A multidimensional meta-analysis of treatments for depression, panic, and generalized anxiety disorder: An empirical examination of the status of empirically supported therapies. *Journal of Consulting and Clinical Psychology*, *69*, 875–899.

Westen, D., Novotny, C. M., & Thompson-Brenner, H. (2004). Empirical status of empirically supported psychotherapies: Assumptions, findings, and reporting in controlled clinical trials. *Psychological Bulletin*, *130*, 631–633.

Westfall, J. M., Mold, M. J., & Fagnan, L. (2007). Practice-based research—"Blue High- ways" on the NIH Roadmap. *Journal of the American Medical Association*, *297*, 403–406.

Received 7 April 2010; revised version received 13 October 2010