Changes in interpersonal problems in the psychotherapeutic treatment of depression as measured by the Inventory of Interpersonal Problems: A systematic review and meta-analysis

Tara McFarquhar\textsuperscript{a,h,*}, Patrick Luyten\textsuperscript{a,h,c,d}, Peter Fonagy\textsuperscript{a,b}

\textsuperscript{a} Research Department of Clinical, Educational and Health Psychology, University College London, London, UK
\textsuperscript{b} Anna Freud National Centre for Children and Families, London, UK
\textsuperscript{c} Faculty of Psychology and Educational Sciences, University of Leuven, Leuven, Belgium
\textsuperscript{d} Yale Child Study Center, Yale University School of Medicine, New Haven, CT, USA

* Correspondence to: Tara McFarquhar, Research Department of Clinical, Educational and Health Psychology, University College London, 1-19 Torrington Place, London WC1E 7HBU. E-mail: tara.mcfarquhar@annafreud.org Telephone: +44 20 7794 2313

**ABBREVIATIONS:** A-CT, Acute-phase Cognitive Therapy; BASIS-24, Behaviour and Symptom Identification Scale; BDI-II, Beck Depression Inventory-II; BSI, Brief Symptom Inventory; CAT, Cognitive Analytic Therapy; CBT, Cognitive-Behavioural Therapy; CBGT, Cognitive-Behavioural Group Therapy; CCT, Client-Centred Therapy; C-CT, Continuation-phase Cognitive Therapy; CT, Cognitive Therapy; EFT, Emotion-Focused Therapy; ES, effect size; HADS, Hospital Anxiety and Depression Scale; HRSD, Hamilton Rating Scale for Depression; IIP-C, Inventory of Interpersonal Problems-Circumplex; IPT, Interpersonal Therapy; ISTDP, Intensive Short-Term Dynamic Program; LTTP, Long-Term Psychodynamic Psychotherapy; MDD, major depressive disorder; MDE, major depressive episode; MHI, Mental Health Index; OQ-45, Outcome Questionnaire 45 item; PA, Psychoanalysis; PCT, Person-Centred Therapy; PET, Process Experiential Therapy; PIT, Psychodynamic-Interpersonal Therapy; PD, Psychodynamic Psychotherapy; PP, Psychoanalytic Psychotherapy; RCT, randomised controlled trial; SCL-90-R, Symptom Checklist-90-Revised; SET, Supportive–Expressive Dynamic Psychotherapy; SFT, Solution-Focused Therapy; STPP, Short-Term Psychodynamic Psychotherapy; TAU, treatment as usual.
ABSTRACT

BACKGROUND: Interpersonal problems are commonly reported by depressed patients, but the effect of psychotherapeutic treatment on them remains unclear. This paper reviews the effectiveness of psychotherapeutic interventions for depression on interpersonal problems as measured by the Inventory of Interpersonal Problems (IIP).

METHODS: An electronic database search identified articles reporting IIP outcome scores for individual adult psychotherapy for depression. A systematic review and, where possible, meta-analysis was conducted.

RESULTS: Twenty-eight studies met inclusion criteria, 10 of which could be included in a meta-analysis investigating changes in the IIP after brief psychotherapy. Reasons for exclusion from the meta-analysis were too few participants with a diagnosis of depression (n=13), IIP means and SDs unobtainable (n=3) and long-term therapy (n=2). A large effect size (g=0.74, 95% CI=0.56–0.93) was found for improvement in IIP scores after brief treatment.

LIMITATIONS: Paucity of IIP reporting and treatment type variability mean results are preliminary. Heterogeneity for improvement in IIP after brief psychotherapy was high (I^2=75%).

CONCLUSIONS: Despite being central to theories of depression, interpersonal problems are infrequently included in outcome studies. Brief psychotherapy was associated with moderate to large effect sizes in reduction in interpersonal problems. Of the dimensions underlying interpersonal behaviour, the dominance dimension may be more amenable to change than the affiliation dimension. Yet, high pre-treatment affiliation appeared to be associated with better outcomes than low affiliation, supporting the theory that more affiliative patients may develop a better therapeutic relationship with the therapist and consequently respond more positively than more hostile patients.

KEY WORDS

IIP; Interpersonal Problems; Depression; Psychotherapy; Systematic Review; Meta-analysis.
INTRODUCTION

Interpersonal problems are both a cause and consequence of depression. They are apparent even in mild depression (Luyten et al. 2005) and are a frequent complaint of those seeking psychotherapy (Horowitz et al. 1988).

Furthermore, several theories have proposed that depression arises when there is a frustration of the basic human need to form and maintain strong and stable relationships (Baumeister and Leary 1995) (see Table 1). This article will examine the evidence to date for changes in interpersonal problems after psychotherapy for depression as assessed with the Inventory for Interpersonal Problems (IIP), one of the most widely used measures to assess interpersonal problems. In addition, we examine whether pre-treatment IIP scores are related to outcome. Where feasible, we present meta-analytic findings.

The IIP (Horowitz et al. 1988) is the only self-report measure of interpersonal problems specifically (as opposed to non-interpersonal problems, e.g. trouble sleeping or eating, unwanted thoughts) and the level of distress caused by them. The IIP originates from interpersonal theories and Horowitz’s interpretation of these approaches in particular.

The theory behind the IIP postulates that behaviours are reciprocally influenced; that is to say, all behaviour invites a reaction. In its original version, the IIP-127 was produced with two sections representing the most frequent ways patients expressed complaints prior to therapy: “It is hard for me to…” and “these are the things I do too much…”. Scoring is on a 5 point Likert scale ranging from 0- not distressed at all by this problem, and 4- extremely distressed by this problem. A shorter factor version of the IIP-127, the IIP-32 (Barkham et al. 1996a) was developed by selecting the items which loaded most highly on their factors and successfully replicated the 8 factor structure with an independent sample. A difference of <.10 in the alpha co-efficients of the scales in the IIP127 and IIP-32 indicated there was acceptable fidelity to the original IIP-127 version. However, it must be noted that the selection of subscale items differs from the other shorter versions.

Despite the apparent robustness of this version of the IIP-32, there is limited use of it in the literature. It seems likely that its demise was due to the increasing popularity of the circumplex versions of the IIP which were published shortly after in 2000 and which included a 64 item and a shorter, 32 item version. The circumplex models of the IIP redress the criticism of the previous factor versions that they fail to inform about the inter-relationship between the scales, resulting only in a list of unrelated factors and were constructed by selecting the 8 ipsatized items that maximised the multiple correlation with each octant identified by a principal component analysis of the IIP127. The IIP-Circumplex (IIP-C) is a 64-item circumplex version (Alden et al. 1990) of the IIP, guided by the interpersonal circumplex model of interpersonal behaviour (Wiggins 1979; Wiggins and Broughton 1985).
It provides a conceptual framework for mapping the relationships between interpersonal problems and is comprised of eight octants of eight items in a circular arrangement, each representing a domain of interpersonal problems, labelled PA, domineering; NO, intrusive; LM, overly nurturant; JK, exploitable; HI, non-assertive; FG, socially avoidant; DE, cold and BC, vindictive. The circumplex can be divided thereafter into quadrants, the top left representing problems associated with a hostile-dominant style, top right a friendly-dominant style, bottom right a friendly-submissive style and bottom left a hostile-submissive style. On the dominance-submission axis (y-axis), behaviours are reciprocal, and on the love-hate axis (x-axis) behaviours are similar, so that, for example, hostile-dominant behaviour solicits a hostile-submissive reaction and friendly dominance a friendly-submissive reaction. Interpersonal problems can be said to be experienced when an individual becomes stuck in a pattern of repeated unwanted and frustrating interpersonal interactions (Horowitz et al. 1997) and the task of the therapist is to break this ‘vicious circle’, first in the therapeutic relationship and later outside of treatment (Horowitz 1996).

Later modifications, including renaming the octants (domineering/controlling, vindictive/self-centred, cold/distant, socially inhibited, non-assertive, overly accommodating, self-sacrificing and intrusive/needy) and producing new normative data, resulted in the IIP-64 (Horowitz et al. 2000). The IIP-64 is composed of the same 64 statements describing common interpersonal problems used in the IIP-C and can be used to evaluate a person’s distress from these problems relative to a standardized sample from the US. Although the ipsatizing methods where different in the development of the IIP-C and the IIP-64, the factor loadings were very similar and the items in each scale are the same, making them easily comparable. It is this 64 item version which is now most commonly used by researchers. Since its manual was published, it has been favoured over the 127 item version except in follow-up studies where the original data was collected with the IIP127 prior to 2000 and where the IIP127 was used in a translated form.

An additional shorter circumplex version, the IIP-SC (Soldz et al. 1995) is a 32 item circumplex version based on the IIP-C (Alden et al. 1990). For each octant, the four items with the highest correlation with the whole scale were selected and verified as reasonable measures of that octant. The correlations between the four item subscales and the eight item subscales of the IIP-C were over $r = 0.9$ and the all the octant scales were within 25º of the expected IIP-C location in a sample of 355 out-patients being treated with psychotherapy, indicating acceptable circumplex properties. The IIP offers several scoring options. The total score, or total mean score, gives an indication of the overall level of distress. Mean subscale scores give a more specific indication of the individual’s interpersonal behaviour in a particular octant. In addition, there are a number of other useful constructs which can be used to describe an individual’s interpersonal behaviour using the IIP. Angular displacement, also referred to as the peak shift or vector angle, indicates the ‘typology’ of the profile (Leary 1957), i.e. the domain in which the individual’s
interpersonal problems mainly lie. **Amplitude** indicates the variability in subscale scores: a score of zero would indicate the same scores on each subscale and a high score a notable peak and trough. **Flux** refers to the amount by which an individual’s interpersonal behaviour varies from the mean on any one specific dimension across social interactions (Moskowitz and Zuroff, 2004). For example, an individual with high flux on the affiliation pole may behave agreeably in one situation and be difficult in another; an individual with low flux would tend to show little variability from their mean level of friendliness regardless of the situation. **Pulse** is the amount of variation in overall extremity of behaviour. It represents the extent to which behaviours vary in their distance from the centre of the interpersonal circle, i.e. how much an individual fluctuates from more extreme to less extreme interpersonal behaviours across events (Moskowitz et al. 2009). Low pulse indicates an individual who continually experiences a similar intensity of feeling (be it high, medium or low); high pulse demonstrates more strong and frequent fluctuations between low- and high-intensity feelings (Kuppens et al. 2007). **Interpersonal spin** is the extent to which an individual’s interpersonal behaviours are dispersed around the circumplex: high spin suggests extensive variation in behaviour and low spin a tendency to repeat similar behaviours over time and across situations (Moskowitz and Zuroff 2004).

Interpersonal problems are the target of many forms of psychotherapeutic treatment, including Dynamic Interpersonal Therapy (DIT), Interpersonal Therapy (IPT) and Emotion-Focused Therapy (EFT), but very little is known about whether they change over the course of treatment, or whether different types of interpersonal problems are associated with outcomes. Two studies are notable for their reporting of interpersonal functioning (not limited to the IIP) over the course of psychotherapy, both specifically Short Term Psychodynamic Psychotherapy (STPP). A meta-analysis investigating the effects of STPP (Abbass et al. 2014), totalling 265 participants across 6 studies with a mean of 15 sessions (range=4-40; SMD -0.42, 95% CI -0.67 to -0.17 in the short-term follow up), reported a small to medium effect size (ES) of treatment on measures of interpersonal functioning including the IIP, the interpersonal relationships subscale of Lambert’s Outcome Questionnaire OQ-45.2 (Lambert 1991) and the interpersonal relations subscale of a scale designed to rate phobic disorders (Alstrom et al. 1983). A meta-analysis of 14 studies of individual STPP (Driessen et al. 2015) reported a large ES (d=0.73, 95% CI 0.50-0.97) for improvement in interpersonal functioning using a number of different measures, including the IIP and subscales of the OQ-45-IR and the Social Adjustment Scale (SAS).

If interpersonal problems per se are trait-like, it would be logical to expect that they would not change, particularly over the course of treatment. However, the distress (state) caused by the problems may change. Therefore, the dimensions of the IIP may not alter very significantly, but the total distress score may reflect change. This has been
demonstrated, for example, over the course of cognitive therapy for depression (Vittengl et al. 2003; Renner et al. 2012) and brief psychodynamic therapy for generalized anxiety disorder (Crits-Christoph et al. 2005). In this latter study, for instance, general distress decreased significantly, whereas the love and dominance dimensions did not, and the mean IIP-C angle (reflecting the most prominent interpersonal style) changed very little from pre- to post-treatment (Crits-Christoph et al. 2005). However, other theories propose that psychotherapy should aim to help the client achieve a less distorted, more balanced interaction between love and dominance by moving them out of their interpersonal routines (e.g. Benjamin, 1974; Wiggins, 1991; Horowitz, 1996; Blatt, 2002; Blatt, 2004; Luyten and Blatt, 2013). Optimum change would result from the client being enabled to modify their behaviour to represent their opposite on the circumplex, for example an individual whose interactions were frequently cold/hostile would be encouraged to move towards more warm/friendly transactions with others (Kiesler, 1983). Considering this, love and dominance scores would be expected to change following successful therapy. Establishing what the literature shows about not only the IIP total score, but also the underlying dimension scores and subscale scores, will be the subject of this review.

With regard to whether (specific) interpersonal problems are related to treatment outcome, studies seem to have yielded inconsistent results and no review to date has attempted to draw conclusions regarding the impact of the IIP octants or quadrants on outcome from any significant number of studies. It has been theorised that the hostile-dominance often present in patients with antisocial features may interfere with the ability to form an alliance in psychotherapy, resulting in a poorer outcome (Gerstley et al. 1989). For instance, in patients with borderline personality disorder, early withdrawal from treatment has been found to be associated with higher levels of pre-treatment hostility (Yeomans et al. 1994; Smith et al. 1995). Patients with a friendly-submissive interpersonal style, by contrast, might be expected to respond more readily to the therapist and to the goals of therapy (hence, treatment will be more successful) than those with a more hostile style.

However, there is very little clarity as to the magnitude of the effect of psychotherapy for depression on IIP scores, or the relationship between pre-treatment IIP and outcome. The purpose of this paper is therefore to conduct a systematic review of studies of the psychotherapeutic treatment of depression that report IIP scores. We focused on the following specific hypotheses: (1), is psychotherapeutic treatment associated with changes between baseline, termination and follow-up in IIP total scores (distress) and in IIP-C dimensions, quadrants or subscale scores, and (2) is there a relationship between therapeutic outcome and IIP total scores (distress), IIP-C dimensions, quadrants or subscale scores. Where there were sufficient data, meta-analyses were performed.

**METHODS**
Search Strategy

An electronic literature search was performed for the years 1946 to December 2016 within PsycInfo and Medline as follows: Inventory of interpersonal problems [title, abstract or methods] OR IIP [all fields]; AND outcome OR response OR improvement OR change [title, abstract]; AND depression OR MDD [all fields]; AND treatment [all fields]. Limo was searched with the following terms: Inventory of interpersonal problems [any field]; AND depression [any field]; AND treatment [any fields]. Reference lists were checked for any additional relevant studies. Papers were included that reported data on the above research questions regarding any psychotherapeutic treatment for depression.

Inclusion criteria were (a) randomised and non-randomised studies reporting IIP outcomes, (b) any version of the IIP, (c) any individual face-to-face psychotherapeutic treatment intervention lasting at least 6 weeks (brief therapy was distinguished from long-term therapy, brief being defined as less than 12 months’ duration and long-term as more than 12 months (Leichsenring et al. 2015)), (d) depression diagnosed by DSM or ICD was the main focus of treatment (psychiatric or medical comorbidities were not excluded), (e) adult participants (18 years or older), (f) inpatient or outpatient setting, (g) published in a peer-review journal between 1946 and 2016, (h) full text available in English.

Exclusion criteria were (a) sample comprising adolescent participants only, (b) sample comprising elderly participants only (one study of patients over 55 years), (c) primary focus of treatment was a disorder other than depression (e.g. eating disorder, anxiety, psychosis, personality disorder), (d) medication-only interventions, (e) group therapy-only interventions, (f) phone/internet-only interventions (g) case studies and series, dissertation abstracts, unpublished theses, books/book chapters/book reviews and letters.

Data extraction and analysis

To conduct a meta-analysis, a minimum of six studies was required (randomised or non-randomised), and means and SDs had to be reported in the original paper or obtained from the authors. The IIP total raw, total global (/mean), total standard T or the global sum scoring methods were all included. In studies where the participants had mixed diagnoses, a cut-off was set at 70% of the sample having a diagnosis of depression (i.e. MDD, Major Depressive Episode (MDE) single or recurrent, double depression, affective disorder) in order to make the sample as homogenous as possible. Studies in which the percentage of the sample diagnosed with depression was <70% were excluded. Separate meta-analyses of pre- to post-treatment scores and post-treatment to follow-up were planned for brief and long-term treatment, to include both IIP total scores and subscales. Potential planned subgroup analyses
focused on potential differences between active psychotherapy vs treatment as usual (TAU)/placebo, and for psychodynamic therapies vs cognitive-behavioural therapy (CBT), as these were the treatments most often reported on in existing studies.

The overall mean ESs were calculated, weighted by the sample size of the individual studies. ESs were converted to Hedges’ g, due to the tendency of Cohen’s d to overestimate ESs in small samples (Borenstein et al. 2009). REVMAN was used to calculate the pooled mean ESs using a random-effects model, which assumes heterogeneity of the included studies (Borenstein et al. 2009). The random-effects model treats studies included as a sample of a population of studies, as opposed to replications of each other, allowing both the random error within studies and the real variations of ES from one study to the next to be accounted for. Accordingly, the results are more conservative, with broader 95% confidence intervals.

The chi-squared test was used to test for heterogeneity: a low P value provides evidence of heterogeneity. The I² statistic was also used to calculate the degree of heterogeneity in percentages, 0% indicating no observed heterogeneity, 25% low, 50% moderate and 75% high heterogeneity (Higgins et al. 2003).

RESULTS

Study Quality

Quality was assessed using a rating system specifically designed for this study, building on the relevant criteria proposed by the Cochrane Collaboration (Higgins and Green 2011). The quality characteristics adopted were similar to those used in Driessen et al’s (2015) meta-analysis of the efficacy of STPP for depression. A six-point scale was adopted, with one point being allocated for each of the following criteria: 1. If more than one treatment arm, participant randomisation (single-arm studies awarded 1 point), 2. 100% of sample have a major depressive disorder (MDD) diagnosis, 3. Manualised treatment, 4. Fully qualified therapists only (students excluded), 5. No concurrent psychotropic medication, 6. Treatment adherence check reported. A score of 0–1 was considered low quality, 2–4 medium quality and 5–6 high quality.

Study Selection

A flow chart of the process of selecting studies for the systematic review is shown in Figure 1. The electronic search yielded 675 potential articles for inclusion. Of these, 122 were found to be duplicates and 414 were excluded following a review of the title and abstract. A full text review was conducted for 139 articles, resulting in 32 articles
to be included. In total, 28 studies (table 2) were included in the systematic review (some were reported across more than one article).

Studies meeting the criteria for inclusion in the meta-analysis are detailed in table 3. Only the meta-analysis of pre-to post-treatment IIP total scores for brief psychotherapy could be performed due to the lack of studies (<6) meeting criteria for IIP outcome of long-term psychotherapy. The planned subgroup analyses were not feasible for the same reason. Short-term active treatment included three psychodynamically oriented samples, four emotion-focused, two cognitive-behavioural, two cognitive, two client-centred, one interpersonal and one solution-focused. The percentage of the sample diagnosed with MDD ranged from 77.5 to 100%. Six of the 10 studies were randomised controlled trials (RCTs).

**IIP change pre-treatment to post-treatment to follow-up**

**Total mean score (Distress)**

There were significant improvements (p values ranged from <0.0001 to <0.05) pre to post active treatment in IIP total mean score for 19 of the 20 studies reporting it. Of these 20, 16 were rated as medium- or high-quality studies. ESs attached to the significant reductions ranged from 0.28 to 2.4 and were large for 11 of the studies. One study reported no significant pre–post change in IIP total scores following STPP (mean treatment period=3.4 months) (Schauenburg et al. 2000).

A meta-analysis of the pre to post ESs for IIP total following brief psychotherapy yielded a large effect (overall pre to post ES g=0.74, 95% CI=0.56–0.93) (see Table 3 and Figure 2). Considerable statistical heterogeneity was observed, \( \tau^2=0.09; \chi^2=55.3, \text{df=14} \) (p<0.00001); \( I^2=75\% \) (Figure 2). A sensitivity analysis was conducted, removing the outlier with an ES >2 (Greenberg & Watson 1998, PET group). This resulted in a slightly decreased overall ES, \( g=0.69, 95\% \text{ CI}=0.52–0.87 \). Heterogeneity remained significant, \( \tau^2=0.07; \chi^2=43.19, \text{df=13} \) (p<0.0001); \( I^2=70\% \).

Six studies reported IIP total score data for a follow-up period. There were not considered suitable for a meta-analysis due to the large variations in treatment length (3–60 months) and follow-up period (1–5 years). However, four of these studies demonstrated a small continued improvement in IIP total score between end of treatment and the follow-up assessment: 1 year post end of a brief Intensive Short-Term Dynamic Program (ISTDP), d=0.22 (Solbakken and Abbass 2015); 1 year post long-term Psychoanalytic Psychotherapy (PP), d=0.17 (Salzer et al. 2010); 3 years post long-term PP, Psychodynamic Psychotherapy (PD) or CBT, d=0.22 (Zimmermann et al. 2015; 5 years
post start of treatment with Solution-Focused Therapy (SFT), STPP or Long-Term Psychodynamic Psychotherapy (LTPP), d=0.16, 0.51 and 0.14, respectively (Lindfors et al. 2015).

Of the studies included in the systematic review, seven were rated as high-quality (24.1%), 12 as medium-quality (41.4%) and 10 as low-quality (34.5%). 72.4% of the studies were randomised or single-arm treatment, 24.1% had a diagnosis of MDD for the full sample, 41.4% adopted a manualised treatment, 27.6% excluded psychotropic medication, 58.6% excluded trainee therapists and 41.4% reported an adherence check of the treatment.

Following the meta-analysis, a sensitivity analysis was conducted to include only those studies rated ‘high quality’ (n=6), resulted in a larger ES (g=0.89, 95% CI=0.65-1.13), and reduced heterogeneity slightly, $\tau^2 = 0.09; \chi^2 = 30.79, \text{df} = 9 (P < 0.0003); I = 71\%$. Excluding the outlier above from the ‘high quality’ analysis, the ES remained large and resulted in lower heterogeneity: g=0.82, 95% CI=0.61-1.03, $\tau^2 = 0.05; \chi^2 = 20.8, \text{df} = 8 (P =0.008); I = 62\%$. The ES for studies rated either ‘low’ or ‘medium’ quality was smaller: g=0.49, 95% CI=0.35-0.63.

A funnel plot (Figure 3) indicated low publication bias once one outlier was removed (Greenberg & Watson 1998, PET group).

**Dimensions (Love/Dominance)**

Four studies included pre to post data for the IIP-C dimensions of love and dominance. Three of these studies reported a significant increase in dominance scores: p<0.01 (Quilty et al. 2013); p<0.01 (Renner et al. 2012); p<0.0001, d=0.32 (Holtforth et al. 2006). One of these studies noted that the increase was no longer significant once pre-treatment elevation had been controlled for (Quilty et al. 2013). One reported no change in dominance (Vittengl et al. 2004). No significant change over the course of treatment was found in love scores in three studies (Vittengl et al. 2004; Renner et al. 2012; Quilty et al. 2013), with a fourth reporting a small significant decrease, p<0.05, d=0.09 (Holtforth et al. 2006).

**Subscales**

Pre to post IIP-C subscale score changes were reported by nine of the studies. Fewer than six studies met criteria for meta-analysis (three brief treatments, one long-term). Seven found significant improvement on all subscales, with the higher-quality studies again tending to report larger ESs (Vittengl et al. 2003; Salzer et al. 2010; Renner et al. 2012; Zimmermann et al. 2015). Large and medium ESs were generally reported more frequently on the affiliative side of the IIP-C. Less change was slightly more notable on the more hostile side, particularly in the hostile-submissive quadrant: one study reported no change on the non-assertive, socially inhibited and the vindictive/self-centred
subscales (Watson et al. 2003), and problems on the socially inhibited and non-assertive subscales were reported as increased by one study (Haase et al. 2008).

**Comparison of modalities**

There were insufficient studies to consider a subgroup analysis of treatment types. However, some results are worthy of note. In the two studies reporting psychotherapy vs. TAU (waitlist) (Bressi et al. 2010; Solbakken and Abbass 2015), the active treatment group had significant reductions in IIP total scores, whereas the TAU groups did not. In the one study reporting IIP total scores for psychotherapy vs. medication (Johansson 2010), the IIP total scores of the psychoanalytic treatment group reduced (d=0.5) more than those of the psychoanalytic group with medication (d=0.32) and the medication-only group (d=0.06).

Due to the wide variability in modalities and high heterogeneity observed, drawing conclusions regarding the benefits of one treatment type over another for interpersonal problems would be premature. However, it is interesting to note the weighting of ESs in the meta-analysis. There does appear to be a slight imbalance of therapeutic approach: the studies arms with large ESs tended to have an emotion-focused therapeutic approach (e.g. Client Centred Therapy (CCT), Interpersonal therapy (IPT), Intensive Short-Term Dynamic Psychotherapy (ISTDP) and Process Experiential Therapy (PET)/EFT), whereas the approach in the studies with smaller ESs leans more towards a goal focus (e.g. CBT, SFT).

**Pre-treatment IIP data and outcome**

Thirteen studies reported data addressing the relationship between pre-treatment IIP scores and outcome. Seven of these met criteria for a meta-analysis, but there were too few studies addressing each pre-treatment IIP construct to perform it.

Of the 13 studies addressing pre-treatment IIP and outcome, eight examined the effect of total IIP distress (only three of which satisfied the meta-analysis criteria). Of these, six studies found an effect of baseline IIP distress, one found no effect and one found a treatment-type dependent effect. Higher pre-treatment IIP distress predicted worse outcome scores over the course of treatment for participants in a study of Cognitive Therapy (CT) (Renner et al. 2012) and in short-term mixed-modality therapy (Ruiz et al. 2004; Beutel et al. 2005). Change in IIP distress was correlated with change in depression scores in Acute phase-Cognitive Therapy (A-CT) (Vittengl et al. 2003; Vittengl et al. 2004) and also in LTPP (Salzer et al. 2010). Participants classed as slow responders to long-term PP or psychoanalysis (PA) had higher levels of pre-treatment distress than faster responders (Berghout et al. 2012). In a
study of CBT, pre-treatment IIP distress were not related to outcome in individual therapy, but in group CBT, more severe IIP distress scores were associated with a poorer outcome (McEvoy et al. 2014). No association was found between pre-treatment distress and the pace of improvement in a study in which patients received PD, CBT or PP (Puschner et al. 2004).

Seven studies reported results for the IIP dimensions of affiliation (love) and dominance, but only four of these satisfied the criterion for meta-analysis of >70% of the sample being diagnosed with depression and only two reported raw scores. All of three studies reported at least one effect, but without much consistency between studies. Higher pre-treatment dominance scores were associated with a poorer outcome in IPT and CBT (Quilty et al. 2013), but predicted a better outcome in CT (Renner et al. 2012), PD (Dinger et al. 2007) and Supportive-Expressive Dynamic Psychotherapy (SET), and no difference in the medication and placebo groups (Dinger et al. 2013). No relationship was found between dominance and outcome in STPP (Schauenburg et al. 2000). Participants with higher pre-treatment affiliation improved more slowly than the low-affiliation participants in a RCT of SET/medication/placebo (Dinger et al. 2013), but higher pre-treatment affiliation was associated with a better outcome in STPP (albeit with small ESs) (Schauenburg et al. 2000). Pre-treatment affiliation did not predict outcome in PD (Dinger et al. 2007). Neither affiliation nor dominance was associated with outcome in PD, CBT and PP treatments analysed together; however, lower affiliation did have a positive effect on outcome for PD alone (Puschner et al. 2004). Higher pre-treatment amplitude (rigidity) was associated with a poorer outcome in IPT and CBT (Quilty et al. 2013) and in short-term mixed modality therapy, even after the effects of total distress were partialed out (Ruiz et al. 2004).

One study reported IIP quadrant data (Puschner et al. 2004) for treatment with either PD, CBT or PP. In the PD group only, the fastest rate of improvement was found for participants with problems in the hostile-submissive quadrant, and the slowest rate of improvement was found for participants with problems in the friendly-dominant quadrant, of the IIP-C. One additional study (Clapp et al. 2014), in which the researchers had had determined their own study-specific IIP-C profiles, reported that 48% of participants with a pre-treatment submissive profile (elevations across the non-assertive, overly accommodating, and self-sacrificing scales) had transitioned to a normative profile at the end of treatment, compared with 57.1% of those with a hostile/withdrawn profile (elevations across the vindictive/self-centred, cold/distant, and socially inhibited subscales).

Three studies investigated the relationship between the IIP subscales and outcome. In PP, the percentage of patients reaching reliable change and clinical significance was highest for participants scoring highly on the overly-accommodating and non-assertive subscales and lowest for the domineering/controlling and cold/distant subscales.
High scores on all the pre-treatment IIP subscales were correlated with a worse outcome in a study of short-term multi-modal treatment, in particular the socially-inhibited and cold/distant subscales (Beutel et al. 2005). However, another study of mixed-modality treatment failed to find correlations between any subscales and outcome (Ruiz et al. 2004).

**DISCUSSION**

The aim of this review was to establish what is known about the effect of psychotherapeutic treatment for depression on interpersonal problems, and whether pre-treatment interpersonal problems relate to therapy outcome. The key finding relates to the change in IIP total or distress score, which was found to improve with a moderate to large ES following brief psychotherapy ($g=0.74$, 95% CI=0.56–0.93) and was consistent with similar findings of improvements in other measures of interpersonal functioning (Driessen et al. 2015; Abbass et al. 2014). Studies rated as ‘high quality’ were associated with the largest ESs. In total, 19 of the 20 studies reporting IIP total scores (distress/) reported a considerable reduction following a course of brief or long-term individual psychotherapy. The majority of these studies (65.5%) were of medium or high quality, and ESs ranged from 0.28 to 2.4. The authors of the one study that found no effect on the IIP total score (Schauenburg et al. 2000) suggest that this was due either to inefficient therapeutic interventions or to the IIP measuring stable ‘traits’, which are unlikely to change to any significant degree following a short-term treatment (in this case 3.4 months). The results of this review contradict the latter argument: it is clear that IIP total scores can change significantly following psychotherapeutic treatments for depression with a duration as short as 4–6 weeks. A more likely explanation for these results would be the participants’ diagnoses: the percentage of those diagnosed with depression was relatively low (23%) compared with the other studies included in this review ($M=64.2\%$, range=20–100%), and adjustment disorder and personality disorder featured more heavily (39% and 32%, respectively). Adjustment disorder might be expected to have a weaker interpersonal component, resulting in less room for change in interpersonal problems, and the treatment period may have to be too short to facilitate change in the relatively high number of participants with personality disorder. The results of the present meta-analysis should be interpreted with caution, due to the high level of heterogeneity observed. This was not unexpected owing to the inevitable need to include several treatment modalities.

A small but notable improvement in IIP total beyond the end of treatment was also found in four of six studies reporting a follow up ($d=0.16–0.51$ for brief treatment and $d=0.14–0.22$ for long-term treatment), a similar effect to that found in Driessen et al’s (2015) meta-analysis of STPP (ES=0.31) for end of treatment to 6 month follow up change. The presence of sleeper effects for interpersonal functioning in these studies suggests that psychotherapy
does seem to have empowered patients to increasingly benefit from positive features in the environment, that is, salutogenesis (Antonovsky 1979).

As to whether pre-treatment IIP total is associated with outcome, there were too few studies for a meta-analysis, and results of individual studies were mixed. Hence, there is still little convincing evidence as to the importance of pre-treatment IIP in predicting the outcome of individual psychotherapy. However, there was an interesting trend towards a greater improvement in IIP total scores in patients receiving EFT relative to those receiving a more ‘goal-focused’ therapy, but at this stage there are insufficient numbers of studies available to investigate ESs for particular treatment modality effects on IIP scores. More consistent reporting of IIP results is becoming more common, and treatment modality effects may be something that can be addressed in the future. The relatively low quality of reporting IIP findings is perhaps surprising given an approximate doubling in citations of ‘Inventory of Interpersonal Problems’ each decade since the 1970s.

Regarding the IIP dimensions (love and dominance), there was some weak evidence to support the theory that love at least is a stable structure (Vittengl et al. 2003). Three of the four studies reporting love found no significant change, and the fourth study found only a very small decrease (d=0.09), however these findings were limited mostly to cognitive therapy and with such few studies, conclusions cannot be drawn as to the stability of love after psychotherapy in general. The stability of dominance was less apparent: only one of the four studies reporting dominance found no change (Vittengl et al. 2004), and the other three studies reported an increase in dominance pre to post treatment. If the treatment modality is considered to affect the change in dominance scores, it does not seem to have been a contributing factor in this case: studies revealed an increase with CT, CBT, IPT and multi-modal therapy. There were no consistent findings regarding the predictive qualities of love or dominance, although three studies did indicate that higher pre-treatment dominance was predictive of better outcomes in CT, PD and SET.

Fairly consistent results were found for the IIP subscales, with the majority of studies reporting significant improvements on all eight scales. Subscales on the affiliative side of the IIP-C seemed generally more amenable to change, which is interesting given the love dimension findings. Perhaps the larger number of studies or the broader range of therapy modalities resulted in a more notable effect. Change in the affiliative subscales seemed also to result in a better outcome than those on the hostile side (particularly the hostile-submissive quadrant), supporting findings that more affiliative people do better in therapy than more hostile ones, perhaps related to the helping alliance (Filak et al. 1986; Gurtman 1996). However, there were insufficient data to draw any meaningful conclusions between the IIP quadrants and outcome: one study reported that patients in the hostile-submissive quadrant undergoing PD improved fastest while those in the friendly-dominant quadrant improved most slowly, while another study reported
that more participants with a hostile profile transitioned to a normative profile at the end of treatment than those with friendly-submissive one.

Two studies presented data to show that higher pre-treatment amplitude (rigidity) was associated with poorer outcomes and that the change in amplitude was associated with change in depression score (Quilty et al. 2013). This could potentially be explained by the ‘epistemic trust’ theory (Fonagy et al. 2015) that more rigid patients are harder to reach in therapy. Unfortunately, there was so little reporting of amplitude scores, along with the other dimensions such as angular displacement, flux, pulse and spin, that no further insight could be gained into their potential to predict response to treatment.

CONCLUSIONS

In conclusion, the meta-analysis revealed that brief psychotherapy was associated with an improvement in IIP total scores. However, the lack of reporting of interpersonal problems in depression trials is starkly apparent. Despite interpersonal problems being the target of many psychotherapeutic treatments, they are not routinely being measured or reported in high-quality trials. A recent Cochrane review of STPP for common mental disorders (Abbass et al. 2014) found that only six of the 33 studies included reported the IIP as an outcome measure. In the present study, we found only 10 randomised studies of psychotherapy for depression that included the IIP as a measure. This study highlights the importance of study quality on changes in pre to post IIP score: higher-quality studies produced an ES almost twice the size of the lower-quality studies. This unexpected finding, (better-quality studies might more usually be expected to have smaller ESs due to their more rigorous design (Cuijpers et al. 2010)) was also noted in the Driessen et al (2015) meta-analysis. Studies that undertook the following quality controls produced larger ESs: randomising participants to treatment arms, diagnosing MDD for the full sample, adopting a manualised treatment, excluding psychotropic medication and trainee therapists, and reporting an adherence check of the treatment. Future investigations of the effects of psychotherapy would be improved by including the IIP-64 as a measure and reporting (i) more dimensional changes, such as total distress, love and dominance, angular displacement, flux, pulse, spin (where repeated measure designs allow) and the subscales; and (ii) the predictive effect of these pre-treatment scores on outcome. Sufficient power to measure these constructs is of course vital. In particular, there is a need for studies comparing therapeutic modalities and comparing psychotherapy with pharmacotherapy to include the IIP, and to report on long-term outcome. A meta-analysis of RCTs of psychotherapy for depression with a 2-year or longer follow-up interval was able to identify only 11 studies (Steinert et al. 2014).

LIMITATIONS
Results were drawn only from journal articles reporting their results in English, meaning there was a risk of failing to include relevant data. An attempt was made to minimise the risk of missing studies by conducting additional manual searches of reference lists and contacting authors directly for data not reported directly in relevant articles. However, there is a risk that this search strategy may still have failed to capture all the data.

The under-reporting of the IIP in studies of psychotherapy for depression meant that a relatively small number of studies could be included in the meta-analysis. Variability in populations and treatment types was a likely contributor to the high level of heterogeneity observed, and made it difficult to draw definite conclusions regarding the predictive value of IIP on outcome.

The assessment of study quality was conducted using a scoring method developed specifically for this study, making comparison to other studies difficult. This decision was taken as the other available tools failed to capture the key quality indicators relevant for this meta-analysis. For example, double-blind rating was not applicable for the IIP as it is a self-report measure. A higher number of patients diagnosed with MDD, medication reporting and treatment adherence checks were considered to be the most important indicators of quality.

In conclusion, this systematic review and meta-analysis reveals that while psychotherapy for depression clearly results in an improvement in interpersonal symptoms at least in brief therapy, the relationship between pre-treatment IIP scores and outcome is still elusive. Dominance does appear to be more amenable to change than affiliation, generally increasing in the limited number of studies reporting it. A rigid interpersonal style may be associated with poorer outcomes, but these speculations await further research.

ACKNOWLEDGEMENTS

The authors wish to thank all investigators who have provided additional information or data regarding their studies upon our request, and Clare Farrar for her assistance with proof reading.

FUNDING

Peter Fonagy is in receipt of a National Institute for Health Research (NIHR) Senior Investigator Award (NF-SI-0514-10157). Peter Fonagy was in part supported by the NIHR Collaboration for Leadership in Applied Health Research and Care (CLAHRC) North Thames at Barts Health NHS Trust. The views expressed are those of the authors and not necessarily those of the NHS, the NIHR or the Department of Health.

This study was supported in part by PhD sponsorship from the Anna Freud National Centre for Children and Families.
REFERENCES


