

**Activity Preferences in Psychotherapy: What Do Patients Want and How Does
This Relate to Outcomes and Alliance?**

Mick Cooper¹
University of Roehampton

Biljana van Rijn
Evi Chryssafidou
Metanoia Institute

William B. Stiles
Miami University, Oxford, USA
and Appalachian State University, Boone, USA

AUTHOR FINAL VERSION

Published as: Cooper, M., van Rijn, B., Chryssafidou, E., & Stiles, W. B. (2021). Activity preferences in psychotherapy: what do patients want and how does this relate to outcomes and alliance? *Counselling Psychology Quarterly*.
<https://doi.org/10.1080/09515070.2021.1877620>

¹ Corresponding author: Mick Cooper, Department of Psychology, University of Roehampton, London SW15 4JD, mick.cooper@roehampton.ac.uk

Activity Preferences in Psychotherapy: What Do Clients Want and How Does This Relate to Outcomes and Alliance?

Abstract

This study aimed to investigate (a) what clients' within-treatment activity preferences were; (b) whether match between preferences and psychotherapy approach predicted outcomes and alliance; (c) whether scores on preference dimensions, per se, predicted outcomes and alliance. Participants were 470 clients engaging in one of five approaches with trainee psychotherapists. We used the Cooper–Norcross Inventory of Preferences to identify clients' within-treatment activity preferences; and multilevel modelling to examine the relationship between these preferences—and match on these preferences—to outcomes and alliance. Clients had an overall preference for therapist directiveness and emotional intensity. We found no evidence of a preference matching effect. Clients who expressed a desire for focused challenge over warm support showed greater progress. Client preferences for focused challenge may be indicative of their readiness to change and indicate a positive prognosis. Further research should directly observe therapeutic practices and assess a range of client variables.

Keywords: aptitude-treatment interaction research, alliance, process research, experiential/existential/humanistic psychotherapy, integrative treatment models

Practical Implications

- Clinicians should be aware that most clients want therapist direction and encouragement to express strong feelings.
- Clients who express a strong preference for warm support over focused challenge may be at particular risk of poor outcomes.
- We did not find evidence that clients did better in therapy if the style of treatment that their clinician was trained in matched their therapy preferences.

Activity Preferences in Psychotherapy: What do Clients Want and How does this Relate to Outcomes and Alliance?

Client preferences are a cornerstone of evidence-based practice (American Psychological Association, 2006). Research indicates that clients are less likely to drop out of therapy, and show better outcomes, if their preferences are accommodated (Swift et al., 2019). Engagement with client preferences has been described as an “ethical imperative” (Norcross & Cooper, 2021). Swift et al. (2019) reported a curvilinear increase in publications using the term “preferences in psychotherapy” from 1970 to 2016, with over 8,000 citations in 2016. A desire to assess and accommodate client preferences is part of a wider movement in the healthcare field towards *personalized medicine* (Norcross & Cooper, 2021). Also known, within psychotherapy, as “treatment adaptation” or “responsiveness”, its aim is to develop and deliver an intervention that is specifically tailored to the unique characteristics, culture, and desires of the client (Norcross & Wampold, 2019).

Swift et al. (Swift et al., 2019) distinguished three types of client preferences. *Activity preferences* refers to “the activities that patients hope they and their therapists will engage in throughout the course of psychotherapy” (p. 157). *Preferences about the therapist* refers to the type of psychotherapist that the client would like to work with. *Treatment preferences* refers to the macro-level intervention approaches that clients might have preferences for, such as psychoanalysis or cognitive behavior therapy.

Whereas most research on activity preferences has focused on preferences regarding the format of intervention (e.g., individual versus group therapy, Renjilian et al., 2001), our research concerned the styles or methods that clients would like their psychotherapists to use. Measures developed to support the investigation of such *within-treatment* activity preferences include the *Psychotherapy Preferences and Experiences Questionnaire* (Sandell et al., 2011), the *Preference for College Counselling Inventory* (Hatchett, 2015), and the *Cooper-Norcross Inventory of Preferences* (C-NIP) (Cooper & Norcross, 2016; Cooper et al., 2019).

Surveys of the general public using the C-NIP—an 18-item inventory with four preference subscales: Therapist Directiveness vs. Client Directiveness (TD–CD), Emotional Intensity vs. Emotional Reserve (EI–ER), Past Orientation vs. Present Orientation (PaO–PrO), and Warm Support vs. Focused Challenge (WS–FC)—have suggested wide variations in within-treatment activity preferences (Cooper & Norcross, 2016; Cooper et al., 2019). However, on average, respondents indicated a preference for Therapist Directiveness over Client Directiveness: with a focus on goals and the acquisition of practical skills. They also indicated, on average, a desire for Emotional Intensity over Emotional Reserve; and no strong preferences on the two other C-NIP subscales. As this research was conducted with members of the general public, however, it is not clear whether these results would generalize to a clinical context. Hence, the first aim of our study (Question 1) was to identify the within-treatment activity preferences of clients at the commencement of therapy. We hypothesized that, as with members of the general public, there would be mean preferences for Therapist Directiveness and Emotional Intensity; but no mean preferences on the two other C-NIP scales: PaO–PrO and WS–FC. At a more exploratory level, we were also interested in whether clients’ activity preferences would vary by demographic characteristics: age, gender, ethnicity, disability status, and sexuality.

Question 2 for this study, central to the practice of personalization, was whether matching to clients’ within-treatment activity preferences would be associated with improved outcomes and alliance. Matching has been hypothesized to have a positive effect on the grounds that clients, at least to some extent, “have a fairly good sense of what they like and what works and does not work for them” (McLeod, 2012, p. 26). That is, an intervention that matches a clients’ preference should provide them with more of what is helpful and less of what is unhelpful (Norcross & Cooper, 2021). However, previous results have been mixed,

and the question remains unsettled, suggesting it may need refinement. In support of this hypothesis, three high quality meta-analyses have indicated that preference assessment and accommodation is associated with improved outcomes (Lindhiem et al., 2014; Swift et al., 2019), a stronger therapeutic alliance (Windle et al., 2019), and reduced dropout (Lindhiem et al., 2014; Swift et al., 2019; Windle et al., 2019). Effects did not vary significantly across type of preference (Swift et al., 2019). However, the majority of studies in these meta-analyses have focused on between-treatment preferences, such as psychotherapy versus medication, rather than activity preferences at the within-treatment level. In the few studies that have examined these more micro-level preferences, no significant effects were found (e.g., Kerns et al., 2014; Kludt & Perlmutter, 1999). In addition, the meta-analyses did not distinguish the effects of treatment matching (*matching effects*) from the effects, on clients, of feeling that they had a choice (*choice effects*), and/or that the psychotherapist was striving to accommodate their preferences (*alliance effects*) (McLeod, 2012; Norcross & Cooper, 2021). Hence, even if an overall positive effect exists for preference assessment and accommodation, it may not mean that matching, per se, is associated with improved outcomes. For this study, the more refined specific hypothesis that we aimed to test was that matching on clients' within-treatment activity preferences would be associated with improved outcomes and alliance.

At an exploratory level, we were also interested in whether clients' within-treatment activity preferences, themselves, might be associated with outcomes (Question 3). For instance, do clients who express a preference for warm support, as opposed to focused challenge, show improved progress in psychotherapy, regardless of what therapy they receive? There is no prior literature on this question and we did not have specific hypotheses here. However, as client factors have been considered the single strongest contributor to outcomes (Bohart & Wade, 2013; Wampold & Imel, 2015), we did wonder if clients' pre-treatment preferences might contribute to outcomes in some way.

Method

Participants

Clients

Clients were recruited at a low-cost counseling and psychotherapy clinic in an urban region of the UK. The service was hosted by a psychotherapy training institution and served as a placement for the institution's trainees. The clients were self-referred, generally following a recommendation by a friend, relative, doctor, or other community provider. Exclusion criteria for the service were severe and enduring mental health problems (such as psychotic disorders or personality disorders), dependent drug or alcohol use as the primary problem, and learning difficulties.

In total, 712 *episodes* of psychotherapy were initiated during our specified time period, May 1st 2016 to August 31st 2018 (an "episode" was defined as a period of ongoing therapy for a client, with no break of longer than three months). As shown in Figure 1, 74 episodes were excluded because individuals did not consent to take part in the research at assessment or subsequently withdrew their consent. We then excluded 31 episodes in which individuals were attending psychotherapy for a second or third time within our inclusion period. Of the remaining 607 individuals assessed for psychotherapy, 21 were excluded because they were not referred for psychotherapy at assessment, and 34 were excluded because they did not start psychotherapy following a referral. We then excluded 62 clients because C-NIP data were not available. A further 12 clients were excluded because of technical problems with their data processing, and 8 because of data entry mistakes (duplicates). Where clients were reallocated to a second psychotherapist, we used only data from the first episode of treatment.

Our final dataset, therefore, consisted of 470 clients who had complete scores on at

least one of the C-NIP subscales, with outcome scores on at least one measure at a minimum of two time points. This sample was predominantly female (67%), White (61.1%), and heterosexual (59.1%), with a mean age of 38.0 years old (Table 1). Overall, 8.3% of participants identified themselves as disabled. Where this was specified it was most commonly chronic pain, sometimes resulting in limited mobility. Approximately two-thirds of the clients were in the clinical range for depression.

Trainee Psychotherapists

The interventions were delivered by 179 trainee psychotherapists, the majority of whom were female ($n = 140$, 78.2%), with a mean age of 42.3 years old ($SD = 9.1$). The median number of clients per therapist was two, with a range of 1–8 clients per therapist.

The trainees were enrolled in Master's or doctoral level programs in one of five different psychotherapy approaches: Gestalt ($n = 7$), humanistic ($n = 50$), integrative ($n = 31$), person-centered ($n = 44$), and transactional analysis ($n = 47$). Prior to commencing their placement, all trainees received one year's didactic training in their approach, and were required to be familiar with their approach's basic concepts and application to practice. Trainees received regular clinical supervision from an experienced psychotherapist at a rate of one hour for every four hours of psychotherapy practice, with supervision meetings at least once every two weeks.

Supervisors

Twenty supervisors were involved in the rating of the therapeutic approaches (TSQ, see below), 11 females and 9 males. The majority ($n = 15$) supervised more than one approach. The minimum qualification for supervisors was a Master's degree in one of the psychotherapies that they supervised. Five of the respondents held a professional doctorate and one was an academic professor. All respondents had a minimum of five years' post-qualification experience.

Measures

We chose a range of outcome measures to capture broad-scale psychological distress, and to match the routine outcomes evaluated in England's Improving Access to Psychological Therapies (IAPT) program. In addition, a trauma outcome measure was used as previous clients at the service where the research was conducted have reported high levels of post-traumatic stress symptoms.

Cooper–Norcross Inventory of Preferences (C–NIP, V.1.0)

The C-NIP assesses respondents' preferences for psychotherapist style (Cooper & Norcross, 2016). This tool, now translated into six languages, was primarily developed for routine assessment of clients' preferences in clinical practice, but has also been used for research (Cooper et al., 2019). The inventory was constructed through principal component analysis of data from a convenience sample of US and UK laypeople and clinicians, responding as prospective clients (Cooper & Norcross, 2016). It is an 18-item instrument that yields scores on four dimensions: TD–CD (5 items), EI–ER (5 items), PaO–PrO (3 items), and WS–FC (5 items) (see Introduction).

The C-NIP invites participants to respond on seven-point, Likert-type items, scaled from -3 to 3 and labeled with opposing preferences (for instance, “Focus on specific goals”—“Not focus on specific goals”). Scale scores equal the unweighted sum of each of the item scores; and hence range from -15 to 15 on the TD-CD, EI-ER, and WS-FC scales, and -9 to 9 on the PaO-PrO scale. In each case, a higher score indicates a preference for the first term in the scale label. The last item on each of the first three subscales is in a reverse direction, as are the first and fourth item on the WS-FC subscale.

In the original measure development study, internal reliabilities were TD-CD Cronbach's $\alpha = .84$, EI-ER $\alpha = .67$, PaO-PrO $\alpha = .73$, and WS-FC $\alpha = .60$ (Cooper & Norcross, 2016). A subsequent online survey with a representative sample of US and UK

members of the general public, again responding as prospective clients, found scale reliabilities of .79, .66, .77, and .55 respectively (Cooper et al., 2019). In the present sample, the scale reliabilities were .67, .54, .61, and .63, respectively. However, on the first three of these scales, these increased to .74, .71, and .80 when the final reversed item was removed. These shortened, internally consistent scales were used in a sensitivity analysis of our data; and the implications of these alpha coefficients are considered in the Discussion.

Cut points for strong preferences in each direction on the four C-NIP scales are TD-CD: $\geq 8, \leq -3$; EI-ER: $\geq 7, \leq -1$; PaO-PrO: $\geq 3, \leq -3$; and WS-FC: $\geq 4, \leq -4$ (Cooper & Norcross, 2016).

Client Health Questionnaire Depression 9-item scale (PHQ-9)

The PHQ-9 is a nine item brief self-report measure for detecting severity of depression symptoms in a general population (Kroenke et al., 2001). Respondents are asked to rate a range of problems over the last two weeks, on a 0 (“not at all”) to 3 (“nearly every day”) scale, indicating the frequency of depressive symptoms. The total scores range from 0 to 27, with higher scores indicating a greater severity of depression, and with a clinical sample cut-off over 9. The PHQ-9 has high internal consistency ($\alpha = 0.89$), good test-retest reliability ($r = .84$), and good convergent validity when correlated with the 20 Item Short Form Survey (SF-20) mental health subscale ($r = .73$) (Kroenke et al., 2001).

Generalized Anxiety Disorder 7-item scale (GAD-7)

The GAD-7 is a brief self-report measure to assess symptom severity of general anxiety disorder (Spitzer et al., 2006). Respondents are asked to rate a range of problems over the last two weeks, on a 0 (“not at all”) to 3 (“nearly every day”) scale, indicating the frequency of anxiety symptoms. The total scores range from 0 to 21, with higher scores indicating a greater severity of anxiety, and with a clinical sample cut-off over 7. The scale has high internal consistency ($\alpha = .92$), high test-retest reliability ($r = .83$), and good convergent validity against the Beck Anxiety Inventory ($r = .72$) (Spitzer et al., 2006).

Clinical Outcomes in Routine Evaluation-10 item version (CORE-10)

The CORE-10 General Distress Measure (CORE Information Management Systems Ltd., 2007) is a short version of CORE-OM (Connell et al., 2007). It includes items on six factors, including anxiety, depression, functioning, risk, trauma, physical symptoms, and risk. Items refer to the previous week and are scored on a 0 (“not at all”) to 4 (“all of the time”) scale. Total scores range for 0 to 40, and scores of 11 or over are in the “clinical” range. The internal reliability of the CORE-10 is high, with $\alpha = .90$; the score for the CORE-10 correlated with the CORE-OM at .94 in a clinical sample and .92 in a non-clinical sample (Barkham et al., 2013).

Work and Social Adjustment Scale (WSAS)

The WSAS (Mundt et al., 2002) is a brief, 5-item self-report scale that measures the respondent’s perceived functional impairment, in relation to an identified problem. Each item rates a dimension of impairment (work, home management, social leisure activities, private leisure activities and relationships) on a 0 (“not at all”) to 8 (“very severely”) scale. Total scores range from 0 to 40, with higher scores indicating more severe impairment, and scores above 10 associated with clinical populations. The measure shows good internal consistency ($\alpha = .82-.93$), test-retest reliability, and clinical predictive validity (Mundt et al., 2002; Thandi et al., 2017); (Zahra et al., 2014).

Impact of Events Scale—Revised (IES-R)

The IES-R is a 22-item self-report measure (for DSM-IV) that assesses subjective distress caused by traumatic events. It is a revised version of the 15-item IES (Horowitz et al., 1979). Items correspond directly to 14 of the 17 DSM-IV symptoms of PTSD. Respondents are asked to refer to a specific stressful event, and, then, rate, for each item, how distressed they have been in the past seven days, using a 0 (“not at all”) to 4 (“extremely”) scale. Total

scores range from 0 to 88. A clinical concern is identified in scores of 24 or more, and a probable diagnosis of PTSD in scores of 33 or above (Weiss & Marmar, 1995). The IES-R shows good internal consistency ($\alpha = .95$) and good convergent validity against measures of PTSD, anxiety, and depression (Beck et al., 2008).

IAPT Phobia Scale (PHO)

The IAPT Phobia Scale is a measure developed as part of the Improving Access to Psychological Therapies (IAPT) program within National Health Service (NHS) settings in the UK, based on the Marks and Matthews (1979) Fear Questionnaire (IAPT, 2011). It consists of three questions, each rating situations that might invoke social anxiety, panic disorder, and specific phobia using a 0 (“would not avoid it”) to 8 (“always avoid it”) scale. A score of four or greater, in each question, is indicative of possible clinical disorder (IAPT, 2011). Data on internal reliability have not been published but in our sample was $\alpha = .75$ at assessment, with convergence against other measures of distress in the range of $r = .37$ (IES-R) to $r = .49$ (WSAS).

Agnew Relationship Measure-5 item version (ARM-5)

The ARM-5 (Cahill et al., 2012) is a 5-item version of the ARM-28, a self-report measure that assesses three dimensions of the therapeutic alliance: bond, partnership, and confidence in the treatment (Agnew-Davies et al., 1998). Respondents are asked to rate a statement that refers to the therapist–client relationship on a 1 (“strongly disagree”) to 7 (“strongly agree”) scale. Summed scores can range from 7 to 35. The internal consistency of the ARM-5 is $\alpha = .79$ (Cahill et al., 2012). For the purposes of our analysis, we used ARM-5 scores at session 10.

Treatment Styles Questionnaire (TSQ)

We designed the TSQ for this study to assess the within-treatment activities that characterized a therapeutic approach from the perspective of experienced professionals/supervisors. The TSQ has four sets of items, corresponding to the four scales on the C-NIP. For the first set of items, respondents were asked to rate on scales from 1 (*strongly disagree*) to 5 (*strongly agree*) how much they thought that clients with a strong preference for “therapist directiveness (focus on goals, structure or techniques)” were “well suited” to each of the five psychotherapeutic approaches used in this study. The three subsequent sets of items asked for similar ratings of suitedness with respect to clients who had strong preferences for emotional intensity, past orientation, and warm support. An initial question asked the respondents to indicate which of the five approaches they supervised.

Intraclass correlation coefficients assessing the reliability of the 19 supervisors who completed ratings in this study (Shrout & Fleiss, 1979, Case 2) indicated excellent rater reliability on the TD-CD (.95), PaO-PrO (.88), and WS-FC (.86) scales; and good reliability on the EI-ER scale (.60) (Fleiss, 1981).

Procedure

Client Recruitment, Assessment, and Selection

Prospective clients contacted the service personally and arranged an assessment appointment. In the majority of cases, clients were informed about the service, or recommended to attend, by their doctors, or by other statutory services such as IAPT. Assesseees were asked to complete the C-NIP, PHQ-9, GAD-7, CORE-10, WSAS, IES-R, and PHO 24 hours prior to assessment, and bring them to their appointment with an assessor, along with their personal demographic form, contact details, and information about their primary physician. Completion of these questionnaires took approximately 15-20 minutes. Five assessors from the clinic participated in assessing prospective clients for eligibility for treatment. The assessment followed established protocol, focusing on clients’ presenting issues, personal histories, treatment histories, health, levels of risk, usage of substances, aims for psychotherapy, preferences about the therapist (gender and ethnicity), and availability.

Where prospective clients met eligibility criteria, the assessor then allocated the client, non-randomly, to a trainee. The trainees worked once a week, with a three-hour slot, throughout their one-year long placement, and could have a maximum of three clients at any given time. Allocation was based primarily on availability: that is, to a trainee who was available for the specific day and time when the client indicated they could attend psychotherapy. Where more than one trainee was available, the assessor prioritized allocation to therapists who had less clients or less completed training hours. In addition, where possible, allocation decisions took clients' gender and ethnicity preferences into account. The clients' C-NIP responses were not used in any way to inform allocation; but information from the assessment, including all C-NIP and other questionnaire data, was available to the trainees. Trainees were not instructed to change their ways of working in response to the questionnaires.

Ethical Approval

Assessors described the clinic's research procedures to clients, responded to questions, and requested consent for participation in research. Clients who declined consent for the research were still offered psychotherapy.

Ethical approval for routine outcomes evaluation was granted on April 21st 2016, with subsequent reviews at each iteration of the protocol. Ethical approval for the dataset of 2018-2019 was granted on the 10th of September 2018 by the respective university ethics committee.

Treatment and Schedule of Measures

Eighteen of the participants were allocated to Gestalt psychotherapy, 107 to humanistic psychotherapy, 86 to integrative psychotherapy, 133 to person-centered psychotherapy, and 134 to transactional analysis.

The clients attended between one and 47 sessions, with a mean of 16.3 sessions ($SD = 11.6$). In 12 cases, where clients, prior to session 5, reported difficulties establishing a therapeutic relationship with their psychotherapist, they were reallocated to a new psychotherapist by their assessor (in these instances, only data from the first allocation was used for our analysis).

Prior to each appointment, clients were asked to complete all outcome measures (GAD-7, PHQ-9, CORE-10, WSAS, IES-R, PHO) and the ARM-5 (from session 2 onwards) and bring the forms to the session. If clients scored below 24 on the IES-R at assessment, the trainees were advised not to administer the measure, although some continued to do so.

Data on dropout (i.e., client withdrawal from treatment before considered advisable by the therapist) were not available.

Therapeutic Approaches

All five programs shared principles and methods consistent with the institution's humanistic, relational philosophy. All trainees were taught to establish a therapeutic alliance based on respect, empathy, and acceptance; listen carefully to clients; facilitate emotional expression; and reflect on their own psychological processes.

Gestalt Psychotherapy. Gestalt therapy is a "process-focused" humanistic psychotherapy which originated with Perls (1951). In their first year training, Gestalt psychotherapy trainees were taught to practice in both non-directive and challenging ways, with a focus on the immediate present, the client's current context ("the field"), and the relationship between psychotherapist and client. They were also taught to explore clients' embodied experiencing.

Person-Centered Psychotherapy. Person-centered psychotherapy, based on the work of Rogers (1951), assumes that distressed individuals have the capacity to address their difficulties if they can explore them with an empathic, supportive, and trustworthy

psychotherapist. In their first-year training, person-centered trainees were taught to practice non-directively, and to work with clients' current problems and circumstances.

Transactional Analysis. Transactional analysis (TA), developed in the 1950s by Eric Berne (Berne, 1958, 1961) combines humanistic and psychoanalytic ideas, conceptualised through the model of "Parent", "Adult", and "Child" ego states. In their first year of training, TA trainees were taught to develop a range of therapeutic agreements with clients; and to work with "ego states" and historical patterns of relating and beliefs ("life scripts"). TA trainees were expected to offer therapeutic direction to clients, to focus on past relationships, and to work on challenging outdated beliefs and experiences.

Humanistic Psychotherapy. Trainees in this approach received combined input on the basic principles of person-centered psychotherapy, Gestalt therapy, and TA. They were encouraged to vary their degree of directiveness, emotional intensity, and focus on past or present depending on the client's presentation and their own emerging style. As with the TA trainees, they were taught to build overt contractual agreements with clients but also to engage in warmly supportive ways.

Integrative Psychotherapy. Integrative training at the institute centered on a model of five different types of therapeutic relating: working alliance, real relationship, transference relationship, reparative relationship, and the transpersonal relationship (Clarkson, 1990). The approach integrated humanistic principles with a psychodynamic, intersubjective emphasis. This psychodynamic emphasis meant that trainees were taught to practice in non-directive ways, but with a focus on past relationships through transference work and reflections on unconscious processes.

Analysis

Question 1: The Within-Treatment Activity Preferences of Clients Embarking on Psychotherapy

Our first question, the within-treatment activity preferences of clients, was examined primarily through descriptive statistics. Differences in C-NIP scale scores across demographic groups were tested using analysis of variance, with Pearson correlations to test associations with age. We adopted an alpha level of $p < .05$ for these tests.

Questions 2 and 3: The Effects of Matching Treatment to Preferences, and of Preferences themselves

The principal dependent variables that we used to analyze Questions 2 and 3 were changes on our outcome measures over the course of psychotherapy. These changes were analyzed using multilevel modelling, with session-by-session outcome scores (level-1) nested within clients (level-2). A three-level model, with clients nested within therapists, was considered. However, this was rejected as variance across therapists on all outcomes and on changes in outcomes over time did not add significantly to model fit. We conducted a single-level regression analysis for therapeutic alliance (ARM-5) scores at session 10. Again, a more complex model was considered—with clients nested within therapists—but was rejected as it did not add significantly to model fit.

As our analysis of the effects of matching (Question 2) aimed to test a specific hypothesis across multiple measures, we used a Bonferroni-corrected alpha of $p < .0071$ ($.05/7$ dependent variables). However, our examination of the effects of preferences, per se (Question 3), was exploratory, and therefore we retained an alpha level of $p < .05$. We also used an alpha of $p < .05$ in our preliminary model development when deciding whether or not to retain demographic and therapeutic approach variables in our models.

Misfit index. To assess the effects of matching treatment to preferences, we developed a *misfit index* for each client. This was the degree of discrepancy between (a) the client's preferences, and (b) their assigned treatment approach, as rated by supervisors on the TSQ. To calculate each client's misfit index, we first standardized clients' scores (across

clients) on each C-NIP scale. Using data from the TSQ, we then also calculated standardized scores for each therapeutic approach on each scale. A raw misfit score was then calculated for each client on each scale, based on the difference between their (standardized) preference, and the (standardized) mean rating of the therapeutic approach that they were allocated to. Next, we squared all our raw misfit scores and took their square roots, so that a misfit in either direction would have a positive value. Finally, we summed the misfit scores across the four scales to yield our overall misfit index.

Dummy Coding of Categorical Variables. Categorical variables were dummy coded as follows: ethnicity (Black and Minority ethnic [BME]/Mixed/other, White; reference category = not stated); disability (disabled, not disabled; reference category = not stated); sexual orientation (gay/lesbian, bisexual; reference category = not stated); and psychotherapeutic approach (Gestalt, TA, humanistic, integrative; reference category = person-centered).

Multilevel Analysis. Procedures for the multilevel analyses followed guidelines proposed by Hox (2010; Hox & Maas, 2005) and Singer and Willet (2003), and were conducted using MLwiN (version 3.02) software with the default iterative generalized least-squares (IGLS) method of estimation. All linear predictor variables, aside from our *TIME* variable, were centered around their grand mean, to ensure interpretability of interaction effects. To examine whether assumptions of normality and linearity had been met, graphs of level-1 and level-2 residuals by rank, and by fixed part predictions, were inspected for the final models (Hox, 2010).

SPSS curve estimation indicated a logarithmic relationship between session number and changes in session scores. Hence we used a natural log transformation of our session numbers as our *TIME* variable.

To establish predictors for each of our outcome scores, we first established unconditional means models. Next, for each outcome, an unconditional growth model was established, which introduced the *TIME* predictor into the model.

To control for individual characteristics, we then entered into the model, as a block, our client-level demographic variables: gender, age, ethnicity, disability status, and sexuality. Each of these independent variables (IVs), along with the $IV \times TIME$ interaction, were tested individually; and all significant predictors within this block were then tested together, with predictors removed if they were no longer significant. With significant predictors retained, we then went on to test a second block of IVs, therapeutic approach, using a similar strategy. Our third block of IVs was client preferences on the four C-NIP scales, tested in a similar way to our two previous blocks (Question 3). Finally, we tested whether misfit indices and, most importantly, the misfit indices $\times TIME$ interactions, would add significantly to model fit (Question 2). To complete the modelling process, a composite model was established, and the contribution of each individual predictor was re-examined, and predictors were removed if they no longer attained significance. Where interactions with *TIME* were significant, direct effects were always retained in the model (Hox, 2010), so that the interactions could be meaningfully interpreted.

Modelling for our therapeutic alliance variable was similar except that interactions with *TIME* were not entered as we focused specifically on ARM-5 scores at session 10.

“Significance” in this modelling process was assessed in two ways. First, through inspection of the predictor’s parameter values against the standard error for this value (the “single parameter test”, Singer & Willett, 2003). Second, through the likelihood ratio statistics test, which compares the deviance statistic (an indicator of model fit) between a model and a more specified version of that model, based on a chi-squared distribution with degrees of freedom equal to the difference in the number of parameters between the models (Hox, 2010).

Sensitivity Analyses. We conducted several sensitivity analyses (using the same alpha cutpoints) to assess the robustness of our findings for Questions 2 and 3. First, we used single-level linear regression modelling on change from baseline to endpoint to see whether a focus on final outcomes, as opposed to session-by-session change, would produce similar results. Second, as matching to strong preferences have been hypothesized to be the key determinant of outcomes (Cooper & Norcross, 2016; Cooper et al., 2019), we conducted 2-way univariate analyses on each of our change and alliance dependent variables to see whether there was an interaction between (a) strong preferences on the C-NIP dimensions, and (b) therapeutic approaches most strongly suited to those preferences (as rated on the TSQ). Two further sensitivity analyses were conducted on our multilevel models to address potential limitations in our design. First, as we were concerned that supervisors who did not practice, or supervise, particular psychotherapeutic approaches might be inaccurate raters of those approaches (on the TSQ), we re-calculated our misfit indices using only the ratings of supervisors of the targeted approaches. Second, as the internal consistency of our C-NIP dimensions was less than optimal, we re-conducted our analyses on the first three C-NIP subscales with re-calculated scores, based only on items that, together, showed internal consistency of $\alpha > .70$.

Results

Preliminary Analysis: Supervisors' Ratings of Psychotherapeutic Approaches

The supervisors' ratings of the five psychotherapeutic approaches on the C-NIP dimensions are presented in Table 2 (with ratings for supervisors of the targeted approaches, only, available online).

On the TD-CD dimension, ratings for all supervisors ranged from 1.6 (person-centered) to 3.6 (TA), on the EI-ER dimension ranged from 3.3 (TA) to 4.1 (Gestalt), on the PaO-PrO dimension ranged from 2.5 (person-centered) to 3.7 (integrative), and on the WS-FC dimension ranged from 3.0 (TA) to 4.1 (person-centered).

C-NIP Descriptives

The mean TD-CD score was 6.6 (95% CI = 6.1–7.0, $SD = 4.6$, range = -12 to 15, $Mdn = 7$, $n = 464$). There were 202 participants (43.5% of all respondents) with strong preferences for TD, and 17 participants (3.7%) for CD. Female clients showed significantly higher preferences for TD than male clients: female mean = 6.9 ($SD = 4.5$, $n = 311$), male mean = 5.9 ($SD = 4.8$, $n = 153$), $F(1, 462) = 4.6$, $p = .03$ ($SMD = 0.21$, 95%CI = 0.02-0.41).

The mean EI-ER score was 5.3 (95% CI = 5.0–5.7, $SD = 4.0$, range = -15 to 15, $Mdn = 5$, $n = 466$). There were 166 participants (35.6%) with strong preferences for EI, and 23 participants (4.9%) for ER. BME clients showed significantly higher preferences for EI than non-BME participants: BME mean = 6.1 ($SD = 4.1$, $n = 129$), non-BME mean = 5.0 ($SD = 4.0$, $n = 337$), $F(1, 464) = 6.9$, $p = .009$ ($SMD = 0.27$, 95%CI = 0.07-0.48).

The mean PaO-PrO score was -0.1 (95% CI = -0.5–0.2, $SD = 3.8$, range = -9 to 9, $Mdn = 0$, $n = 464$). There were 107 participants (23.0%) with strong preferences for PaO, and 114 participants (24.6%) for PrO. Clients with a disability showed significantly higher preferences for PaO than clients who did not indicate that they had a disability: disabled mean = 1.2 ($SD = 5.1$, $n = 39$), no disability or not stated mean = -0.3 ($SD = 3.7$, $n = 425$), $F(1, 462) = 5.0$, $p = .03$ ($SMD = 0.39$, 95%CI = 0.06-0.72).

The mean WS-FC score was -1.8 (95% CI = -2.2–-1.4, $SD = 4.8$, range = -15 to 15, $Mdn = -1$, $n = 458$). There were 43 participants (9.4%) with strong preferences for WS, and 151 participants (33.0%) for FC. Female clients showed significantly higher preferences for WS than male clients: female mean = -1.3 ($SD = 4.8$, $n = 305$), male mean = -2.8 ($SD = 4.8$, $n = 153$), $F(1, 456) = 9.5$, $p = .002$ ($SMD = 0.31$, 95%CI = 0.12-0.51). In addition, bisexual clients showed significantly higher preferences for warm support than non-bisexual clients:

bisexual mean = 1.9 ($SD = 5.5$, $n = 23$), non-bisexual mean = -2.0 ($SD = 4.7$, $n = 435$), $F(1, 456) = 14.3$, $p < .001$ ($SMD = 0.82$, $95\%CI = 0.40-1.24$).

Model Development

Final models for our seven outcomes variables are presented in Table 3. Visual inspection of residuals suggested criteria for normality and linearity had been met.

Preliminary Model Development

All outcome measure scores showed significant improvements over time, with both intercepts and improvements over time showing greatest contributions to model fit when allowed to vary randomly at the client level.

With respect to demographic characteristics, our most consistent finding was that disabled clients had higher overall scores on all measures of distress than non-disabled clients. This ranged from 14.42 points higher on the IES-R (for clients coded as “disabled” versus those not coded as “disabled”) to 2.23 points higher on the WSAS (for clients not coded as “not disabled” versus those coded as “not disabled”). In addition, on five of the six outcome measures, White clients had lower levels of overall distress than BME clients. This ranged from 4.60 points lower on the IES-R to 1.24 points lower on the GAD-7 (for clients coded as “White” versus those not coded as “White”). However, there was no evidence that disability status or ethnicity was related to changes in outcomes over time. Clients coded as “heterosexual” showed greater reductions in depression (PHQ-9) and anxiety (GAD-7) over time, as compared with those not coded as “heterosexual”; and improvements over time were less for clients coded as “bisexual” on the IES-R. On three of the outcome measures, clients coded as “heterosexual” also showed higher overall scores. Males showed lower levels of anxiety than females on the GAD-7 ($b = -1.36$), but also poorer alliances at session 10 on the ARM-5 ($b = -1.36$). Older age was associated with lower overall scores on the WSAS ($b = -0.08$), but also with less reduction over time ($b = 0.02$).

Therapeutic approach did not contribute significantly to fit on any of our seven models.

Association between Preferences and Outcomes (Question 3)

Higher scores on the WS-FC dimension were significantly associated with lower improvements over time on four of the six outcome measures, ranging from $b = 0.04$ (PHO) to $b = 0.08$ (CORE-10). This means that clients who indicated a preference for warm support at assessment tended to improve less than those who indicated a preference for focused challenge. Suppose, for instance, that two clients both scored 13 on the PHQ-9 at assessment (the sample mean), but one scored 10 on the WS-FC dimension and the other scored -10. Holding all other variables constant, our model would suggest that, by session 20, the client preferring warm support would have improved by 4.3 points, while the client preferring focused challenge would have improved by 7.9 points.

Two further sets of findings suggested that clients who wanted more active input from their psychotherapists tended to improve more over time. First, on the GAD-7 ($b = -0.05$) and the PHO ($b = -0.04$), a greater desire for therapist directiveness (TD-CD) was associated with more improvements. Second, on the WSAS, a greater desire for emotional intensity (EI-ER) was associated with more improvement over time ($b = -0.09$).

A greater desire for therapist directiveness (TD-CD) was associated with higher levels of overall distress on five of our outcome measures, ranging from $b = 0.15$ (PHQ-9) to $b = 0.57$ (IES-R).

Association between Misfit and Outcomes (Question 2)

In the analysis of our main hypothesis (Question 2), total misfit between preferences and allocated therapeutic approach was not associated with change on any of our outcome measures over time. Single parameter scores (b/SE) ranged from 0.42 to -0.68. Total misfit

was also not significantly associated with therapeutic alliance on the ARM-5, $b = .21$, $SE = 0.11$.

Sensitivity Analyses

Results from our separate linear regression models were substantially the same as from our multilevel analyses (see online material). Misfit was not associated with change on any outcomes; and clients who preferred focused challenge over warm support showed significantly greater change on the PHQ-9 ($b = -0.11$), CORE-10 ($b = -0.19$), and PHO ($b = -0.11$). In our categorical analyses, we did not find any evidence that clients with strong preferences on the C-NIP dimensions did better in the therapeutic approaches matched to those preferences than clients with strong preferences in the opposite directions. No significant matching effects were found when we used only the TSQ ratings from supervisors of the targeted therapeutic approaches. We also found no significant matching effects on the outcome measures when we used the shortened C-NIP scales with internal consistency $> .7$. However, when only these scores were used, we found a significant positive association between misfit and alliance on the ARM-5, $b = .40$, $SE = .14$, $p = .002$ (indicating that more misfit was associated with a greater alliance at session 10). The latter sensitivity analyses also showed no substantial deviations from the relations between preferences and outcomes shown in Table 3.

Discussion

Perhaps the most novel finding from our study (in answer to Question 3) was that clients who expressed a preference for focused challenge (and other, more active therapist activities) tended to do better in psychotherapy than those who expressed a preference for warm support. One possible explanation for these findings is that clients' preferences for more active therapist inputs are indications of their motivation, engagement, or readiness to change: client factors that are known to be amongst the most important determinants of improvement (Bohart & Wade, 2013). Clients who express a preference for challenge and confrontation, for instance, may be more likely to be in the "preparation" or "action" "stages of change" (Krebs et al., 2018; Prochaska & DiClemente, 1986), while those wanting warm support may be still be in the "contemplation" or "precontemplation" stages. A related interpretation is that a client's desire for warm support could indicate a hope, or expectation, of therapist sympathy, rather than a drive towards personal transformation. It might also reflect a greater dependency on the therapist.

Another possible explanation is that clients who desire focused challenge may tend to enter psychotherapy with problems that are specific and solvable, or with change that is measurable on standard outcome tools; whereas clients who desire warm support may tend to enter psychotherapy with more amorphous and difficult-to-measure problems. Such differences may reflect clients' problems being at different stages of assimilation (e.g., Stiles, 2011). Unassimilated problems are experienced as amorphous, and the process of bringing them into clearer awareness may not immediately reduce distress, even though it is therapeutically necessary, according to the assimilation model (Basto et al., 2017). Supportive, rather than challenging, therapeutic approaches may be a better fit for facilitating this uncovering process (Stiles et al., 1992; Stiles et al., 1997).

Our findings here, however, are limited by the low internal consistency of the C-NIP subscales (in particular the WS-FC dimension), and by the relatively narrow range of therapies (humanistic/relational) within which these effects were identified. Future research would benefit from studying the relationship between preferences and change in other therapeutic approaches such as CBT. Development of more reliable means of assessing within-treatment activity preference is also essential for further studies.

In answer to our Question 2, and in contrast to the recent meta-analyses (Lindhiem et al., 2014; Swift et al., 2019; Windle et al., 2019), we found no evidence of a preference

matching effect. (Indeed, the one nominally significant finding, from our sensitivity analyses, indicated that greater misfit was associated with a stronger therapeutic alliance.) Our null findings concur, however, with the more closely similar previous studies (e.g., Kerns et al., 2014; Kludt & Perlmutter, 1999) in suggesting that matching on within-treatment activity preferences does not lead to differentially positive benefits. As one way to understand this, psychotherapists and clients may have been involved in an ongoing, interactive, and implicit process of adjustment and responsiveness to the client's preferences within the framework of the therapist's approach, with both parties playing an active and agentic role. To the extent that this responsive accommodation was successful, each client may have received an individually optimized treatment, overcoming any effects of mismatching (Stiles et al., 1998). In effect, the responsiveness may have tended to minimize the actual levels of misfit. In addition to their clinical sensitivity, the trainees had access to the clients' C-NIP scores, (though it is unclear how much they examined them), and they may have tailored their work to fit the clients' preferences.

There are also several important limitations to our study design that could explain our null result, even in the absence of responsiveness effects. First, our assessment was based on ratings of named therapeutic approaches rather than direct observations of practices, and our TSQ measure was novel and has not been separately validated. Second, all of the treatments were delivered by early-stage trainees so their results may not be representative of these treatments as usually practiced. Third, our range of therapeutic approaches was relatively narrow: all based on relational and humanistic principles and some major approaches unrepresented (for instance, no CBT). Fourth, as noted earlier, the internal consistency of our preference scales was limited. Fifth, we did not assess dropout, which has been found to have the clearest association with preference accommodation to date (Swift et al., 2018). And sixth, allocation to therapeutic approach was not random.

Non-significant findings may also have come about because we looked at preference matching effects across all clients, rather than specifying the particular clients for whom matching effects may have been most impactful. Norcross and Cooper (2021), for instance, hypothesized that matching effects are likely to be greatest when individuals have high levels of *motive congruence*: where their explicitly recognized wants and needs match their implicit wants and needs (Thrash et al., 2012). Such motive congruence has been found to be higher in self-determining individuals, and in people who are more sensitive to their bodily states and less monitoring of other's expectations (Thrash et al., 2012). In addition, motive congruence is closely associated with *domain familiarity*: how knowledgeable people are about a particular context (Lichtenstein & Slovic, 2006). Hence, matching effects might be hypothesized to be greatest in clients who have previous knowledge, and experience, of psychotherapy; while those new to the domain might be predicted to have less understanding of what will be most suited to them.

Subsequent studies of within-treatment activity preference matching effects, therefore, might profitably (a) directly assess therapeutic practices, using (b) validated instruments, with interventions delivered by (c) experienced psychotherapists, who are (d) randomly allocated and (e) providing a wide diversity of therapeutic approaches. In addition, there should be assessment of (f) dropout and (g) client variables, in particular previous experience of psychotherapy.

In answer to our Question 1, the distributions of activity preferences of clients in our UK clinical sample showed only modest differences from that of the representative sample of UK and US general public (Cooper et al., 2019). On the TD-CD subscale, around 40% in each sample showed a strong preference for therapist directiveness, with less than 5% showing a strong preference for client directiveness. On the EI-ER subscale, we found that a higher percentage of our clinical sample showed a strong preference for emotional intensity

(35.9%) as compared with the general public sample (20.0%). On the PaO-PrO dimension, the present sample showed a more even split between strong preferences for past and present orientation (22.7% and 24.8%, respectively) as compared with the general public sample (19.7% and 38.4%, respectively). On the WS-FC dimension, we found a higher percentage of participants showing strong preferences for focused challenge (32.6% vs 9.4% for warm support), as compared with the general public sample (23.2% vs 25.4% for warm support). Thus, these results converge with findings from previous studies to provide robust evidence that people, whether members of the general public or clients about to embark on psychotherapy, tend to want direction from a psychotherapist in the form of skills, goals, and structure; and encouragement to express their emotions.

Two of our significant demographic predictors of C-NIP scores were consistent with previous findings (Cooper et al., 2017). First, BME clients showed a higher preference for emotional intensity than non-BME clients. It is not clear why this is the case. Indeed, some BME, collectivist cultures place greater emphasis and valuing on low emotional arousal as compared with more individualist, Western cultures (Lim, 2016). Second, females showed a greater preference for warm support over focused challenge, as compared with males. This finding would seem to be consistent with social role theories of sex differences (e.g., Eagly & Wood, 2016), by which men are expected to act in more dominant, agentic, and independent ways (Koenig, 2018).

As with all our analyses, these findings are limited by the low internal consistency of the C-NIP scales. A priority for further research, therefore, is to ensure that activity preferences are measured using indicators of proven reliability. The exclusion of certain clients from our sample also means that these findings cannot be generalized to the full population of clients—in particular, those with more severe and enduring mental health problems. Nevertheless, given the importance of tailoring treatment to the specific needs and wants of particular cultures, genders, and other demographic groupings, this inquiry into preference differences across demographic characteristics is an essential area for further enquiry.

Among this study's implications for practice, psychotherapists should be aware that most—though not all—clients come to treatment wanting direction from their psychotherapists and encouragement to express strong feelings. They should also be aware that female clients may want more warm support than males, and that BME clients may have a particular preference for emotional intensity. Our study did not provide evidence that outcomes and alliance are improved if psychotherapists accommodate such preferences; but other findings (Lindhiem et al., 2014; Swift et al., 2019; Windle et al., 2019)—as well as ethical consideration, such as the desire to support client autonomy—suggest that psychotherapists should be cognizant of such tendencies. On the other hand, we also found considerable diversity across clients' preferences, supporting a view that psychotherapists should never assume what individual client's preferences are, but engage in a process of preference assessment and dialogue (Norcross & Cooper, 2021; Schmid, 2002). In assessing client preferences, either through measures or through dialogue, it may be important to take note when clients are expressing strong preferences for support over challenge or are in other ways indicating a desire for a less active psychotherapist stance. Our findings suggest that such clients may be at particular risk of poor outcomes. Talking to clients about the meaning of such preferences may be important, particularly to ascertain if it is linked to poor motivation for treatment or a precontemplation stage of change.

Conclusion

Our study is the first to systematically explore the activity preferences of clients at the commencement of psychotherapy. Consistent with previous findings from general public samples, we found that clients generally prefer more directive and emotionally intense

activities; with females showing a greater preference for warm support than males, and BME clients showing a greater preference for emotional intensity than non-BME clients. We did not find evidence of a preference matching effect, though methodological limitations in our study make any conclusions here premature. Of greatest importance, perhaps, we found an association between a preference for focused challenge and improved psychotherapy outcomes. This result is consistent with research on client factors in psychotherapy and, whilst requiring further empirical support and exploration, may have important implications for the assessment of readiness for—and likelihood of—change.

References

- Agnew-Davies, R., Stiles, W. B., Hardy, G. E., Barkham, M., & Shapiro, D. A. (1998). Alliance structure assessed by the Agnew Relationship Measure (ARM). *British Journal of Clinical Psychology, 37*(2), 155-172.
- American Psychological Association. (2006). Evidence-based practice in psychology. *American Psychologist, 61*(4), 271-285.
- Barkham, M., Bewick, B., Mullin, T., Gilbody, S., Connell, J., Cahill, J., Mellor-Clark, J., Richards, D., Unsworth, G., & Evans, C. (2013). The CORE-10: A short measure of psychological distress for routine use in the psychological therapies. *Counselling and Psychotherapy Research, 13*(1), 3-13. <https://doi.org/10.1080/14733145.2012.729069>
- Basto, I., Pinheiro, P., Stiles, W. B., Rijo, D., & Salgado, J. (2017). Changes in symptom intensity and emotion valence during the process of assimilation of a problematic experience: A quantitative study of a good outcome case of cognitive-behavioral therapy. *Psychotherapy Research, 27*(4), 437-449. <https://doi.org/10.1080/10503307.2015.1119325>
- Beck, J. G., Grant, D. M., Read, J. P., Clapp, J. D., Coffey, S. F., Miller, L. M., & Palyo, S. A. (2008). The Impact of Event Scale-Revised: Psychometric properties in a sample of motor vehicle accident survivors. *Journal of anxiety disorders, 22*(2), 187-198.
- Berne, E. (1958). Transactional analysis: a new and effective method of group therapy. *American Journal of Psychotherapy, 12*(4), 293-309.
- Berne, E. (1961). *Transactional Analysis in Psychotherapy*. New York: Grove Press.
- Bohart, A. C., & Wade, A. G. (2013). The client in psychotherapy. In M. J. Lambert (Ed.), *Bergin and Garfield's Handbook of Psychotherapy and Behavior Change* (6th ed., pp. 219-257). Chicago: John Wiley and Sons.
- Cahill, J., Stiles, W. B., Barkham, M., Hardy, G. E., Stone, G., Agnew-Davies, R., & Unsworth, G. (2012). Two short forms of the Agnew Relationship Measure: The ARM-5 and ARM-12. *Psychotherapy Research, 22*(3), 241-255.
- Clarkson, P. (1990). A Multiplicity of Psychotherapeutic Relationships. *British Journal of Psychotherapy, 7*, 148-163.
- Connell, J., Barkham, M., Stiles, W. B., Twigg, E., Singleton, N., Evans, O., & Miles, J. N. V. (2007). Distribution of CORE-OM scores in a general population, clinical cut-off points and comparison with the CIS-R. *British Journal of Psychiatry, 190*, 69-74.
- Cooper, M., & Norcross, J. C. (2016). A Brief, Multidimensional Measure of Clients' Therapy Preferences: The Cooper-Norcross Inventory of Preferences (C-NIP). *International Journal of Clinical and Health Psychology, 16*(1), 87-98. <https://doi.org/10.1016/j.ijchp.2015.08.003>
- Cooper, M., Norcross, J. C., Raymond-Barker, B., & Hogan, T. P. (2017). [Unpublished dataset].
- Cooper, M., Norcross, J. C., Raymond-Barker, B., & Hogan, T. P. (2019). Psychotherapy preferences of laypersons and mental health professionals: Whose therapy is it? *Psychotherapy, 56*, 205-216. <https://doi.org/10.1037/pst0000226>

- Eagly, A. H., & Wood, W. (2016). Social role theory of sex differences. *The Wiley Blackwell encyclopedia of gender and sexuality studies*, 1-3.
- Fleiss, J. L. (1981). Balanced incomplete block designs for inter-rater reliability studies. *Applied Psychological Measurement*, 5(1), 105-112.
- Hatchett, G. T. (2015). Development of the Preferences for College Counseling Inventory. *Journal of College Counseling*, 18(1), 37-48. <https://doi.org/10.1002/j.2161-1882.2015.00067.x>
- Horowitz, M., Wilner, N., & Alvarez, W. (1979). Impact of Event Scale: A measure of subjective stress. *Psychosomatic medicine*, 41(3), 209-218.
- IAPT. (2011). *The IAPT data handbook: Guidance on recording and monitoring outcomes to support local evidence-based practice* (2.0.1 ed.). London: IAPT.
- Kerns, R. D., Burns, J. W., Shulman, M., Jensen, M. P., Nielson, W. R., Czapinski, R., Dallas, M. I., Chatkoff, D., Sellinger, J., & Heapy, A. (2014). Can we improve cognitive-behavioral therapy for chronic back pain treatment engagement and adherence? A controlled trial of tailored versus standard therapy. *Health Psychology*, 33(9), 938-947. <https://doi.org/10.1037/a0034406>
- Kludt, C. J., & Perlmutter, L. (1999). Effects of control and motivation on treatment outcome. *Journal of Psychoactive Drugs*, 31(4), 405-414.
- Koenig, A. M. (2018). Comparing Prescriptive and Descriptive Gender Stereotypes About Children, Adults, and the Elderly. [Original Research]. *Frontiers in Psychology*, 9(1086). <https://doi.org/10.3389/fpsyg.2018.01086>
- Krebs, P., Norcross, J. C., Nicholson, J. M., & Prochaska, J. O. (2018). Stages of change and psychotherapy outcomes: A review and meta-analysis. *Journal of Clinical Psychology*, 74(11), 1964-1979. <https://doi.org/10.1002/jclp.22683>
- Kroenke, K., Spitzer, R. L., & Williams, J. B. (2001). The PHQ-9. *Journal of General Internal Medicine*, 16(9), 606-613. <https://doi.org/10.1046/j.1525-1497.2001.016009606.x>
- Lim, N. (2016). Cultural differences in emotion: differences in emotional arousal level between the East and the West. *Integrative medicine research*, 5(2), 105-109. <https://doi.org/10.1016/j.imr.2016.03.004>
- Lindhiem, O., Bennett, C. B., Trentacosta, C. J., & McLear, C. (2014). Client preferences affect treatment satisfaction, completion, and clinical outcome: A meta-analysis. *Clinical Psychology Review*, 34(6), 506-517. <https://doi.org/10.1016/j.cpr.2014.06.002>
- Marks, I. M., & Mathews, A. M. (1979). Brief standard self-rating for phobic patients. *Behaviour Research and Therapy*, 17(3), 263-267.
- McLeod, J. (2012). What do clients want from therapy? A practice-friendly review of research into client preferences. *European Journal of Psychotherapy and Counselling*, 14(1), 19-32.
- Mundt, J. C., Marks, I. M., Shear, M. K., & Greist, J. M. (2002). The Work and Social Adjustment Scale: a simple measure of impairment in functioning. *The British Journal of Psychiatry*, 180(5), 461-464. <https://doi.org/10.1192/bjp.180.5.461>
- Norcross, J. C., & Cooper, M. (2021). *Personalizing psychotherapy: Assessing and accommodating client preferences*. Washington: APA.
- Norcross, J. C., & Wampold, B. E. (2019). Evidence-based psychotherapy responsiveness: The third task force. In J. C. Norcross (Ed.), *Psychotherapy relationships that work* (3rd ed., Vol. 2: Evidence-based therapist responsiveness, pp. 1-14). NY: Oxford University.
- Perls, F., Hefferline, R. F., & Goodman, P. (1951). *Gestalt Therapy: Excitement and Growth in the Human Personality*. New York: Julian Press.

- Prochaska, J. O., & DiClemente, C. C. (1986). Toward a comprehensive model of change. In W. R. Miller & N. Heather (Eds.), *Treating addictive behaviors* (pp. 3-27): Springer.
- Renjilian, D. A., Perri, M. G., Nezu, A. M., McKelvey, W. F., Shermer, R. L., & Anton, S. D. (2001). Individual versus group therapy for obesity: effects of matching participants to their treatment preferences. *Journal of Consulting and Clinical Psychology, 69*(4), 717-721. <https://doi.org/10.1037/0022-006X.69.4.717>
- Rogers, C. R. (1951). *Client-Centered Therapy*. Boston: Houghton and Mifflin.
- Sandell, R., Clinton, D., Frövenholt, J., & Bragesjö, M. (2011). Credibility clusters, preferences, and helpfulness beliefs for specific forms of psychotherapy. *Psychology and Psychotherapy: Theory, Research and Practice, 84*(4), 425-441. <https://doi.org/10.1111/j.2044-8341.2010.02010.x>
- Schmid, P. F. (2002). Knowledge of acknowledgement? Psychotherapy as 'the art of not-knowing' – prospects on further developments of a radical paradigm. *Person-Centered and Experiential Psychotherapies, 1*(1&2), 56-70.
- Shrout, P. E., & Fleiss, J. L. (1979). Intraclass correlations: uses in assessing rater reliability. *Psychological Bulletin, 86*(2), 420-428.
- Spitzer, R. L., Kroenke, K., Williams, J. B., & Löwe, B. (2006). A brief measure for assessing generalized anxiety disorder: The GAD-7. *Archives of Internal Medicine, 166*(10), 1092-1097. <https://doi.org/10.1001/archinte.166.10.1092>
- Stiles, W., Barkham, M., Shapiro, D., & Firth-Cozens, J. (1992). Treatment order and thematic continuity between contrasting psychotherapies: Exploring an implication of the assimilation model. *Psychotherapy Research, 2*(2), 112-124.
- Stiles, W. B. (2011). Coming to terms. *Psychotherapy Research, 21*, 367-384. <https://doi.org/10.1080/10503307.2011.582186>
- Stiles, W. B., Honos-Webb, L., & Surko, M. (1998). Responsiveness in psychotherapy. *Clinical Psychology-Science and Practice, 5*(4), 439-458.
- Stiles, W. B., Shankland, M. C., Wright, J., & Field, S. D. (1997). Aptitude-treatment interactions based on clients' assimilation of their presenting problems. *Journal of Consulting and Clinical Psychology, 65*(5), 889-893.
- Swift, J. K., Callahan, J. L., Cooper, M., & Parkin, S. R. (2018). The impact of accommodation client preferences in psychotherapy: A meta-analysis. *Journal of Clinical Psychology, 74*(11), 1924-1937. <https://doi.org/10.1002/jclp.22680>
- Swift, J. K., Callahan, J. L., Cooper, M., & Parkin, S. R. (2019). Preferences. In J. C. Norcross (Ed.), *Psychotherapy relationships that work* (3rd ed., pp. 157-187). NY: Oxford University.
- Thandi, G., Fear, N. T., & Chalder, T. (2017). A comparison of the work and social adjustment scale (WSAS) across different patient populations using Rasch analysis and exploratory factor analysis. *Journal of psychosomatic research, 92*, 45-48. <https://doi.org/10.1016/j.jpsychores.2016.11.009>
- Thrash, T. M., Maruskin, L. A., & Martin, C. C. (2012). Implicit-explicit motive congruence. In R. M. Ryan (Ed.), *The Oxford handbook of human motivation* (pp. 141-156). New York: Oxford University Press.
- Wampold, B. E., & Imel, Z. E. (2015). *The Great Psychotherapy Debate: The Evidence for What Makes Psychotherapy Work* (2nd ed.). New York: Routledge.
- Weiss, D. S., & Marmar, C. R. (1995). The impact of event scale--revised. In J. P. Wilson & T. M. Kean (Eds.), *Assessing psychological trauma and PTSD: A practitioner's handbook* (pp. 399-411). NY: Guilford.
- Windle, E., Tee, H., Sabitova, A., Jovanovic, N., Priebe, S., & Carr, C. (2019). Association of Patient Treatment Preference With Dropout and Clinical Outcomes in Adult

Psychosocial Mental Health Interventions: A Systematic Review and Meta-analysis.

JAMA Psychiatry. <https://doi.org/10.1001/jamapsychiatry.2019.3750>

Zahra, D., Qureshi, A., Henley, W., Taylor, R., Quinn, C., Pooler, J., Hardy, G., Newbold, A., & Byng, R. (2014). The work and social adjustment scale: Reliability, sensitivity and value. *International Journal of Psychiatry in Clinical Practice*, 18(2), 131-138. <https://doi.org/10.3109/13651501.2014.894072>

Table 1.*Participant Demographics and Service Use Data (N = 470)*

Age (Mean, SD)	38.0 (12.3)
Gender (n, %)	
Female	315 (67.0%)
Male	155 (33.0%)
Ethnicity (n, %)	
Asian	47 (10.0%)
Black	29 (6.2%)
Mixed	38 (8.1%)
White	287 (61.1%)
Other	16 (3.4%)
Not stated	53 (11.3%)
Disability (n, %)	
Yes	39 (8.3%)
No	316 (67.2%)
Not stated	115 (24.5%)
Sexual orientation (n, %)	
Gay/lesbian	27 (5.7%)
Bisexual	23 (4.9%)
Heterosexual	278 (59.1%)
Not stated	142 (30.2%)
Severity of depression at assessment (PHQ-9) (n, %)	
Minimal or none (0-4)	36 (7.7%)
Mild (5-9)	107 (22.8%)
Moderate (10-14)	121 (25.7%)
Moderately severe (15-19)	97 (20.6%)
Severe (20-27)	83 (17.7%)
Missing data	26 (5.5%)
N sessions (Mean, SD)	16.2 (11.5)
Engagement (n, %)	
Early termination (1-4 sessions)	94 (20.0%)
Engagers (5+ sessions)	376 (80.0%)
Reallocation (n, %)	
Reallocated	11 (2.3%)
Not reallocated	459 (97.7 %)
Psychotherapy orientation (n, %)	
Gestalt	18 (3.8%)
Humanistic	105 (22.3%)
Integrative	85 (18.1%)
Person-Centered	132 (28.1%)
TA	130 (27.7 .0%)

Table 2.*Treatment Styles Questionnaire Scale Means and Standard Deviations, by Supervisors'**Theoretical Approach*

Scale	Gestalt <i>Mean (SD, n)</i>	Humanistic <i>Mean (SD, n)</i>	Integrative <i>Mean (SD, n)</i>	Person- Centered <i>Mean (SD, n)</i>	TA <i>Mean (SD, n)</i>
All raters					
Therapist directiveness (TD-CD)	2.8 (0.8, 19)	2.7 (0.8, 19)	2.9 (0.8, 19)	1.6 (0.8, 19)	3.6 (0.8, 20)
Emotional intensity (EI-ER)	4.1 (1.0, 19)	3.7 (0.9, 19)	3.5 (0.9, 19)	3.4 (1.3, 19)	3.3 (1.0, 20)
Past orientation (PaO-PrO)	2.8 (0.8, 19)	3.0 (0.7, 19)	3.7 (0.7, 19)	2.5 (0.8, 19)	3.7 (0.8, 20)
Warm support (WS-FC)	3.0 (0.9, 19)	3.6 (0.8, 19)	3.1 (0.7, 19)	4.1 (1.1, 19)	3.0 (0.8, 20)

Note. TD-CD = Therapist Directiveness vs. Client Directiveness, EI-ER = Emotional Intensity vs. Emotional Reserve, PaO-PrO = Past Orientation vs. Present Orientation, WS-FC = Warm Support vs. Focused Challenge. Mean scores indicate average ratings (5 = strong agreement, 1 = strong disagreement) of the suitedness of the psychotherapeutic approach to patients with a strong preference for the first (left hand) term in the C-NIP dimension. Pairs of cells in bold indicate psychotherapeutic approaches at opposite ends of each dimension used for the categorical analysis.

Table 3.

	PHQ-9 (<i>n</i> = 452)	GAD-7 (<i>n</i> = 463)	CORE-10 (<i>n</i> = 453)	WSAS (<i>n</i> = 454)	IES-R (<i>n</i> = 421)	PHO (<i>n</i> = 450)	ARM5 (<i>n</i> = 261)
Intercept	13.15 (0.54) + u_{0j}	12.67 (0.40) + u_{0j}	20.39 (0.52) + u_{0j}	21.11 (1.48) + u_{0j}	45.986 (1.38) + u_{0j}	7.49 (0.41) + u_{0j}	33.22 (0.26)
<i>TIME</i>	-1.19 (0.16) + u_{1j}	-1.36 (0.09) + u_{1j}	-2.12 (0.13) + u_{1j}	-1.77 (0.48) + u_{1j}	-5.61 (0.44) + u_{1j}	-0.41 (0.08) + u_{1j}	
Demographics							
Gender		Male: -1.38 (0.45)					Male: -1.38 (0.44)
Gender* <i>TIME</i>							
Age				-0.08 (0.03)			
Age* <i>TIME</i>				0.02 (0.01)			
Ethnicity	White: -1.29 (0.53)	White: -1.24 (0.44)	White: -1.86 (0.62)	BME: 1.80 (0.86)	White: -4.60 (1.62)		
Ethnicity* <i>TIME</i>							
Disability	Disabled: 3.38 (0.91)	Disabled: 2.44 (0.75)	Disabled 3.64 (1.09)	Not Disabled: -2.23 (0.83)	Disabled: 14.42 (2.77)	Disabled: 2.71 (0.86)	
Disability* <i>TIME</i>							
Sexuality	[Heterosexual: 0.09 (0.61)]			[Heterosexual: 1.19 (0.85)]	[Bisexual: -6.76 (4.01)]	Heterosexual: -1.36 (0.50); Homosexual: -2.52 (1.07)	
Sexuality* <i>TIME</i>	Heterosexual* <i>TIME</i> : -0.44 (0.21)			Heterosexual* <i>TIME</i> : -0.80 (0.29)	Bisexual* <i>TIME</i> : 4.55 (1.86)		
Therapeutic Approach							
Gestalt							
Gestalt* <i>TIME</i>							
Humanistic							
Humanistic* <i>TIME</i>							
Integrative							
Integrative* <i>TIME</i>							
Person-centered							
Person-centered* <i>TIME</i>							
TA							
TA* <i>TIME</i>							
Preferences							
TD-CD	0.15 (0.06)	0.30 (0.05)	0.29 (0.07)		0.57 (0.17)	0.17 (0.06)	
TD-CD* <i>TIME</i>		-0.05 (0.02)				-0.04 (0.02)	
EI-ER				[0.13 (0.11)]			
EI-ER* <i>TIME</i>				-0.09 (0.04)			
PaO-PrO							
PaO-PrO* <i>TIME</i>							
WS-FC	[-0.03 (0.06)]		[0.01 (0.07)]	[0.01 (0.09)]		[0.06 (0.05)]	
WS-FC* <i>TIME</i>	0.06 (0.02)		0.08 (0.03)	0.06 (0.03)		0.04 (0.02)	
Misfit							
Total misfit						0.36 (0.14)	
Total misfit* <i>TIME</i>							

Regression coefficients and standard errors (*b*, *SE*) in final models

Note. u_{0j} = random variance for intercept, u_{1j} = random variance for *TIME*. Square brackets indicate non-significant direct effects entered to make significant interactions interpretable. Higher numbers on all outcome scores indicate greater distress. Higher numbers on ARM-5 indicate greater alliance.

Figure 1. *Participant Flow Chart*