Does Client-Therapist Gender Matching Influence Therapy Course or Outcome in Psychotherapy?

Michael J Lambert*

Brigham Young University-Hawaii, Provo, Utah, USA

Abstract

The present study examined the impact of therapist and client gender, and gender matching on therapy outcome and attendance variables. Analysis was based on a database of over 17,000 students treated in a university counseling center by over 200 therapists. Results indicated that gender of clients (but not the gender of therapist, or the match between therapists and clients) was associated with greater improvement in female clients. Clinical significance analyses indicate that female clients were more likely to start treatment in the clinical range and also end treatment in the “improved” category. Significant associations were found between client gender, therapist gender, and gender match on the total number of sessions attended by clients with females receiving more sessions.

Keywords: Psychotherapy outcome; Psychotherapy and gender; Gender effects; Attrition; OQ-45; Treatment effects; Clinically significant change

Introduction

One of the earliest reviews of the effects of client-therapist gender matching on therapy outcomes was conducted by Berzins [1]. He concluded that the limited research available at the time suggested that gender is a weak contributor to the outcome of therapy. However, he encouraged further examination of the influence of gender matching in therapy as, “its obvious relation to sex-role expectations and stereotypes in clinical settings makes it an important variable for empirical reexamination” (p. 232).

Reviews of the research that has been conducted since that decade continue to echo earlier conclusions – that the effects of gender matching on therapy outcome are weak at best. A review by Mogul [2] concluded that “numerical research on patient populations regarding the effects of therapist gender has offered little of definitive or predictive value” (p. 9; see also Sterling et al. [3]). Bowman [4] reviewed the literature up to that point and concluded that “the view that therapist sex is a poor predictor of outcome in therapy is the most conservative and probably the most sensible position” (p. 684). More recently, Beutler et al. [5] concluded that the relationship between therapist sex and outcome has been even less consistent in contemporary research when compared to studies of earlier decades.

Although the majority of studies point to the lack of effect of gender matching on therapy outcomes, therapists continue to have an array of clinical opinions, of varying intensity regarding the utility and importance of gender matching. More specifically, clinicians appear divided as to whether mixed or same gender matching is likely to lead to better psychotherapy outcomes or whether gender matching is a relevant consideration for influencing therapy outcomes at all [4,6]. Berzins [1] observation that sex-role stereotypes and expectations influence clinical opinion appears to be a topic of contemporary as well as historical relevance.

One example of gender based assumptions influencing clinical recommendations is found in sexual abuse treatment literature. Some professionals have recommended that females who have been sexually abused be treated by female therapists [2,6]. These recommendations appear based on clinical assumptions rather than research findings (for example “male therapists may quite inadvertently victimise incest survivors due to their own enculturation...”); ([7], p. 81). In a series of studies that explored this clinical recommendation, it was found that female children aged between 7 and 15 years who were sexually abused demonstrated no differences in level of comfort with a male or female therapist following the initial appointment, and while the majority of girls indicated an initial preference for a female therapist a significant number did not (25%; [8]). Furthermore, there were no differences in comfort with therapist based on therapist gender at the conclusion of brief-term therapy [9] and in a randomized study there were no differences in therapy outcomes for sexually abused girls dependent on whether they saw a male or female therapist [10]. While this series of studies does not exclude the possibility that there may be some benefits to gender matching in sexual abuse treatment for some clients in some therapy situations, they do suggest that the question of gender matching is a complex one, not easily amenable to broad recommendations.

Other clinical populations have also been examined to determine if gender matching influences therapy outcomes. One of the most widely cited examinations of the influence of gender on therapy outcomes was conducted by Zlotnic et al. [11]. They observed that previous research looking at gender matching was limited due to use of relatively small samples, the use of female clients only, the lack of valid and reliable outcome measures, and non-equivalent distribution of clients in terms of symptom severity or diagnoses. Using data from the NIMH treatment of Depression Collaborative Research Program, Zlotnic et al. [11] examined the influence of the following variables on treatment outcomes for 203 participants: a) therapist gender (i.e., do male or female therapists have better outcomes), b) gender matching, and c) client preference for gender matching with therapist gender. These variables had no statistically significant impact on change in Hamilton...
Rating Scale for Depression scores, Empathy ratings, or Attrition. One weakness of the study was relatively low power to detect potential effects with the regression analysis used. The authors called for further studies with larger samples. There appears to have been minimal response to this recommendation.

The most recent meta-analysis to explore the influence of gender and gender matching in therapy, combined data from 64 studies [12]. The authors concluded “the effect sizes for female and male clients were not significantly different than zero, suggesting no advantage for female or male therapists when seeing clients of the same or opposite sex. Taken together with the relatively low overall effect size, these data indicated that there was essentially no difference in the effect associated with therapist sex” (p. 145).

Although there are few studies that support the notion that gender is a relevant consideration to therapy outcome in a global sense, there are some instances of specific findings for some client subgroups. For example, Whaley assessed the effects of gender matching in a sample of 124 African-American male patients participating with paranoia in the context of “severe mental illness.” Participants were interviewed by a race matched male or female therapist for an intake interview – those in the gender matched condition reported less paranoid symptoms, but more cultural mistrust. In another study, exploring response to drug abuse treatment - female clients, Latino clients, and older clients who were gender matched with therapists had slightly higher rates of abstinence when compared to gender-mismatched pairings [13].

Gender effects have also been identified for variables that relate to the course of therapy. For example, Sue et al. [14] examined data from Automated Information System (AIS) maintained by the Los Angeles County Department of Mental Health; from which they drew a sample of over 13,000 clients classified as Asian American, African American, Mexican American or White. Their analysis found that gender matching increased the likelihood that clients returned after the initial session. This effect held for clients classified as Asian American or White but not for the other ethnic groups. Gender matched clients were also more likely to have greater treatment length. This effect was apparent in the White and Mexican American groups only. However, gender matching was not related to treatment outcome (defined as change in DSM-III, Global Assessment of Functioning Scores) for any of the ethnic groupings.

In conclusion, there appears to be very limited contemporary or historical research that supports the notion that positive outcome in therapy is enhanced by gender matching. Where relationships are observed in gender matching research they primarily relate to indexes of likelihood of dropout, or duration of therapy. However, this overall pattern of findings is often not reflected in treatment literature recommendations or decision making in clinical practice – where some in the field continue to advocate for gender matching. The current study was designed to address some of the limitations of previous studies in this area and expand the breadth of information available for clinical decision making by exploring the effects of gender and gender matching in the context of a college counseling center. The study was designed to address the following three questions: Does therapy outcome differ based on the either the gender of the client, the gender of the therapist, or the gender match between the client and the therapist? Does the duration of therapy (total number of sessions) differ based on the gender of the client, the gender of the therapist or the gender match between the client and the therapist? Does the non-return after one session rate differ based on the gender match between the client and the therapist?

### Methods

#### Participants

Clients: The client sample for this study consisted of college students seen at a large university counseling service for individual psychotherapy between 1996 and 2008. Treatment was available at no charge to fulltime students of the university. Clients at the center presented with a wide range of problems from simple homesickness to personality disorders. There were no session limits imposed. Data has been collected since 1996 as part of the clinical routine. Prior to each therapy appointment (including intake) clients complete a 45 item measure of symptom distress (Outcome Questionnaire–45; [15]). Data used in the present study represents all clients presenting for individual therapy between 1997 and 2008 that had valid OQ-45 data in the database.

Sample one: This sample was used in evaluating therapy outcome (question one in introduction) and consisted of students who saw only one therapist at the clinic and completed at least two sessions of counseling (so that change scores could be calculated). The number of students in this sample was 6,628 (4,078 female, 2,550 male). The average age was 22.7 years (SD 4.14). The range of sessions completed in this sample was 2 to 93 with a mean of 5.53.

Sample two: This sample was used to evaluate the duration of therapy (question two in introduction) and consisted of students who saw only one therapist at the clinic and completed one or more sessions of counseling. The number of students that completed at least one session of counseling was 10,746 (6,292 female, 4,454 male). The average age was 22.8 years (SD 4.19). The range of sessions completed in this sample was 1 to 93, with a mean of 3.79 (SD 5.40).

Sample three: This sample was used to evaluate the rate of non-return after one session of treatment. Unlike samples one and two it included students that switched therapists during the course of treatment – either after the first session or subsequently. The number of students in this sample was 17,340 (10,370 female, 6,970 male). The average age was 22.6 years (SD 3.94). The range of sessions completed in this sample was 1 to 211, with a mean of 6.90 (SD 9.76).

Therapists: Two-hundred and eighteen therapists (86 female, 124 male) contributed data to the entire data pool of 10,746 clients. Therapists varied on level of training (preinternship, internship, and postinternship), type of training (clinical psychology, counseling psychology, social work, marriage and family therapy), and primary theoretical orientation (cognitive–behavioral, behavioral, humanistic, psychodynamic). The modal therapist was a male, licensed, counseling psychologist with a doctorate, who identified his primary theoretical orientation as cognitive–behavioral. Procedures for case allocation varied over the study period but were dominated by allocating cases based on counselor availability rather than employing systematic methods to match therapists with specific client variables.

#### Measure

Client progress in this study was tracked using the Outcome Questionnaire (OQ-45), a 45-item self-report measure developed specifically for the purpose of tracking and assessing client outcomes in a therapeutic setting. The OQ-45 is a well-established instrument that has been validated across the country and across a broad range of normal and client populations. Lambert et al. [15] reported an internal consistency for the OQ-45 of .93 and a 3-week test–retest value of .84 both of which are considered adequate. Concurrent validity figures
were calculated by comparing the OQ-45 total score with total scores from other measures including the Symptom Checklist-90 [16], Beck Depression Inventory [17], Zung Depression Scale [18] and the State-Trait Anxiety Inventory [19]. All of the concurrent validity figures with the OQ-45 and these instruments were significant at the .01 level with a range of r's from 0.50 to 0.85. Most important, the OQ-45 has been shown to be sensitive to the effects of interventions on client functioning [20,21].

The OQ-45 is scored using a 5-point scale (0-never, 1-rarely, 2-sometimes, 3-frequently, 4-almost always), which yields a possible range of scores from 0 to 180. High scores on the OQ-45 indicate more distress and as clients improve scores decrease. Although not used in this study, the OQ-45 has three subscales that measure quality of interpersonal relations, social role functioning, and symptom distress. The total score, which provides a global assessment of functioning, was used in this study. Using formulas developed by Jacobson and Truax [22], clinical and normative data for the OQ-45 were analyzed by Lambert et al. [15] to provide cutoff scores for the Reliable Change Index (RCI). Clients who change in a positive or negative direction by at least 14 points are regarded as having made "reliable change." Clinically significant change, as defined by Jacobson and Truax [22], also involves moving from a score typical of a dysfunctional population to a score typical of a functional one. The cutoff on the OQ-45 for marking the point at which a person's score is more likely to come from the dysfunctional population than a functional population has been estimated to be 64. When a client's score falls at, or below, 63 it is concluded that this client's functioning is similar to a non-client's level of functioning at that point in time.

Clients who show reliable change and pass the clinical cutoff into the normal range are considered "recovered"; those who only show reliable change are considered "improved." Clients who do not change more than 14 points in a positive or negative direction are considered "no change," and clients who worsen by 14 points are considered "deteriorated." Support for the validity of the OQ-45's reliable change and clinical significance cutoff scores have been reported by Lunnen and Ogles [23] and Bauer et al. [24].

**Results**

The respective therapy variables for Samples 1 through 3 are summarized in Tables 1-3. Separate two-way ANOVA's on these samples indicated a main effect for client gender on initial OQ-45 score (Sample 1: F1,6624=64.65, p<0.001, Sample 2: F1,10742=210.05, p<0.001, Sample 3: F1,17336=308.48, p<0.001). Effects for therapist gender and client X therapist gender interaction were non-significant. The magnitude of the difference in initial OQ-45 score on client gender for samples 1 through 3 was 4.95, 6.98, and 6.58 points respectively.

### Table 1: Initial client OQ-45 total score and change in OQ-45 total score for sample 1: All clients with two sessions or more with same therapist.

<table>
<thead>
<tr>
<th>Group</th>
<th>N</th>
<th>Initial OQ-45 score</th>
<th>Change OQ-45 score</th>
</tr>
</thead>
<tbody>
<tr>
<td>All clients</td>
<td></td>
<td>Mean</td>
<td>SD</td>
</tr>
<tr>
<td>Male Therapists</td>
<td>2496</td>
<td>69.82</td>
<td>22.69</td>
</tr>
<tr>
<td>All Female Therapists</td>
<td>4132</td>
<td>67.86</td>
<td>22.50</td>
</tr>
<tr>
<td>All Male Clients</td>
<td>2550</td>
<td>65.38</td>
<td>23.23</td>
</tr>
<tr>
<td>All Female Clients</td>
<td>4087</td>
<td>70.62</td>
<td>21.93</td>
</tr>
<tr>
<td>Male-Male</td>
<td>1880</td>
<td>65.02</td>
<td>22.98</td>
</tr>
<tr>
<td>Male-Female</td>
<td>670</td>
<td>66.39</td>
<td>23.91</td>
</tr>
<tr>
<td>Female-Male</td>
<td>2252</td>
<td>70.24</td>
<td>21.80</td>
</tr>
<tr>
<td>Female-Female</td>
<td>1826</td>
<td>71.08</td>
<td>22.09</td>
</tr>
</tbody>
</table>

The chi-squared analyses were also conducted to determine if there were significant differences between the proportions of clients who were classified as recovered, improved, no change and deteriorated in each of the matching categories (see above for operationalization of these categories). A further analysis using the same techniques considered only those clients whose initial score was in the clinical range and the proportion of these clients who were classified as recovered, improved, no change, and deteriorated at the end of treatment.

A chi-squared analysis using a contingency table approach was used to address the following question: Does the non-return after one session rate differ based on the gender match between the client and the therapist? For all significant chi-squared analyses the Marascuillo procedure was employed to compare probability differences between all possible pairs of proportions. The Marascuillo procedure allows for calculation of a unique critical value for each pairing that takes into account the number of pairings to be compared, the sample size and the proportions that are being contrasted according to the following formula:

\[
z = \frac{\hat{p}1 - \hat{p}2}{\sqrt{\frac{1}{n1} + \frac{1}{n2}}},
\]

Where:

- \(z\) = the critical value for the comparison of two proportions \(p_1\) and \(p_2\)
- \(n\) = the critical value for the chi-squared distribution at user defined alpha level for \(k\)-1 degrees of freedom where \(k\) is the number of proportions that are being contrasted.
- \(n\) = sample size for group \(i\)

The proportion of clients who started in the clinical range for each of the four client-therapist gender pairings. Chi-squared analyses were also conducted to determine if there were significant differences between the proportions of clients who were classified as recovered, improved, no change and deteriorated in each of the matching categories (see above for operationalization of these categories). A further analysis using the same techniques considered only those clients whose initial score was in the clinical range and the proportion of these clients who were classified as recovered, improved, no change, and deteriorated at the end of treatment.

A chi-squared analysis using a contingency table approach was used to address the following question: Does the non-return after one session rate differ based on the gender match between the client and the therapist? For all significant chi-squared analyses the Marascuillo procedure was employed to compare probability differences between all possible pairs of proportions. The Marascuillo procedure allows for calculation of a unique critical value for each pairing that takes into account the number of pairings to be compared, the sample size and the proportions that are being contrasted according to the following formula:

\[
z = \frac{\hat{p}1 - \hat{p}2}{\sqrt{\frac{1}{n1} + \frac{1}{n2}}},
\]

Where:

- \(z\) = the critical value for the comparison of two proportions \(p_1\) and \(p_2\)
- \(n\) = the critical value for the chi-squared distribution at user defined alpha level for \(k\)-1 degrees of freedom where \(k\) is the number of proportions that are being contrasted.
- \(n\) = sample size for group \(i\)

The respective therapy variables for Samples 1 through 3 are summarized in Tables 1-3. Separate two-way ANOVA's on these samples indicated a main effect for client gender on initial OQ-45 score (Sample 1: F1,6624=64.65, p<0.001, Sample 2: F1,10742=210.05, p<0.001, Sample 3: F1,17336=308.48, p<0.001). Effects for therapist gender and client X therapist gender interaction were non-significant. The magnitude of the difference in initial OQ-45 score on client gender for samples 1 through 3 was 4.95, 6.98, and 6.58 points respectively.
The analysis for the entire sample 1 identified significant differences between client-therapist groupings for the proportion of clients who were categorized as starting treatment in the clinical range, ending treatment in the normal range, improved by the end of treatment, and recovered by the end of treatment. The Marascuillo procedure indicated that where significant differences could be identified between client-therapist gender groupings, the differences reflected differences between male and female clients, with female clients tending to demonstrate a higher likelihood of starting treatment in the clinical range and a higher likelihood of being in the improved category. There were no significant results suggesting that outcomes for male or female clients were different based on whether they were gender matched or mismatched with therapists. Considering only sample 1 clients who started in the clinical range, the chi-squared analysis summarized in Table 5 indicated significant differences between client-therapist groupings for the proportion of clients who were categorized as improved by the end of treatment, and deteriorated by the end of treatment. No significant differences between individual client-therapist pairings were identified by the Marascuillo procedure.

**Does therapy outcome differ based on the gender of the client, the gender of the therapist or the gender match between the client and the therapist?**

A two-way ANOVA was employed with client gender and therapist gender as factors and total number of sessions as the dependent variable; initial OQ-45 Score was used as a covariate – see Table 6. At alpha=0.05, the results indicated a statistically significant main effect for client gender when controlling for initial OQ score (on average female clients stayed in therapy 0.58 sessions longer than male clients); significant effects were also found for therapist gender (on average clients of male therapists remained in therapy 0.2 sessions longer than clients of female therapists). Due to the significant effect of initial OQ score, this variable was entered as a covariate in the model for all subsequent ANOVA analyses.

**Clinical significance analysis**

The proportion of clients in sample one who started in the clinical range, and the various outcome categories of clients are summarized in Table 5, as are the outcomes for those clients who started treatment in the clinical range.

The analysis for the entire sample 1 identified significant differences between client-therapist groupings for the proportion of clients who...
clients of female therapists); there was a significant interaction effect (see Figure 1) which is dominated by the following features: Female clients tend to remain in therapy longer than male clients, clients of male therapists tend to remain in therapy longer, and the increase in average number of sessions for male clients if they are seen by a male therapist as opposed to a female therapist (d=0.12) is larger than the effect seen for female clients (d=0.02).

Thus in answer to question two above – our analysis indicates that female clients tend to have significantly (statistically) longer duration of therapy compared to male clients (d=0.11). The difference in number of sessions between male and female clients is 0.58 of a session. Clients of male therapists have a significantly longer duration of therapy compared to female clients – effect size d=0.04. Both female and male clients are likely to have a longer duration of therapy if seen by a male therapist (0.10 and 0.63 sessions longer respectively). Despite being statistically significant, the strength of both of these effects is classified as small (d=0.02 and d=0.12 respectively for females and males) (Figure 2).

**Does the 'non-return after one session' rate differ based on the gender match between the client and the therapist?**

The proportion of clients who only attended a single session of treatment is summarized in Table 3. Using a contingency table approach we tested the null hypothesis that proportion of single session clients for the four client-therapist combinations do not significantly differ from one another. The results indicate that there were significant differences between the groups, χ²=85.65, p<0.001. The percentage of clients with single sessions is represented in Figure 3.

The Marascuillo procedure was used to compare all possible proportion pairs to identify which displayed significant differences, the results are presented in Table 7. The results indicate that while males had a significantly higher single session rate than female clients (27.3% vs. 21.4%), these rates did not significantly differ based on the gender of the therapist that they saw.

**Discussion**

The present study examined the impact of therapist and client gender, and gender matching on therapy outcome and process.

**Table 5: Clinical significance category proportions for each client-therapist combination for all sample 1 clients, and sample 1 clients whose initial score was in the clinical range.**

<table>
<thead>
<tr>
<th>Source of Variation</th>
<th>Degrees of Freedom</th>
<th>Sums of Squares</th>
<th>Mean Square</th>
<th>F Statistic</th>
<th>p Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Therapist Gender</td>
<td>1</td>
<td>329.65</td>
<td>329.51</td>
<td>11.50</td>
<td>0.001</td>
</tr>
<tr>
<td>Client Gender</td>
<td>1</td>
<td>605.76</td>
<td>605.76</td>
<td>21.14</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Interaction</td>
<td>1</td>
<td>158.70</td>
<td>158.70</td>
<td>5.54</td>
<td>0.019</td>
</tr>
<tr>
<td>Error</td>
<td>10741</td>
<td>307717.59</td>
<td>28.65</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>10746</td>
<td>467963.00</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Table 6: Two way ANOVA: effects of client gender and therapist gender on duration of therapy.**

![Figure 2: Interaction of client gender and therapist gender on total number of sessions – (means displayed are corrected for initial OQ-45 score).](image)
variables. Analysis was based on a database of over 17,000 students treated in a university counseling center by over 200 therapists. The results of this study indicate that the gender of clients (but not the gender of therapist, or the match between therapists and clients) has a statistically significant effect on the magnitude of therapeutic improvement as measured by the OQ-45; favoring female clients (2.61 points, \(d=0.14\)). Clinical significance analysis indicates that female clients are more likely to start treatment in the clinical range and also more likely to be in the “improved” category; there were no differential effects identified based on gender matching variables.

Furthermore, significant associations were found between the client gender, therapist gender, and gender match on the total number of sessions attended by clients. On average, female clients tend to remain in therapy longer than male clients by 0.58 sessions), clients of male therapists tend to remain in therapy longer by 0.2 sessions, and a significant interaction effect was dominated by the observation that the increase in average number of sessions for male clients if they are seen by a male therapist as opposed to a female therapist is larger than the effect seen for female clients; average increase of 0.64 and 0.12 sessions for male and female clients respectively when corrected for initial OQ-45 total score.

Finally, male clients tend to have a higher likelihood that they will only attend one session of therapy than female clients (27.3% vs. 21.4%) but these rates were not significantly impacted by the gender of the therapist that they first met with.

The statistically significant results found in this study highlight the need for incorporating additional statistical methods to interpret findings from studies with large sample sizes. With large enough sample size even negligible differences between groups on a pragmatic level can reach statistical significance. By incorporating other statistical measures, such as effect size and clinical significance indices, the implications of statistically significant findings could be clarified.

In terms of differences in outcomes and duration of therapy significant effects were found but the effect sizes were small (range 0.02<-d<-0.14) and thus the practical implications of these findings on clinical decision making would appear to be minimal. The overall conclusion is that gender of the client or therapist, or the gender match between the client and the therapist has negligible impact on therapy outcome or duration of therapy for the average client. From a fiscal point of view, the results may have some relevance – based on the results of treatment duration and outcome, if a clinic sees 1000 new male clients a year it would cost them 640 more sessions if these clients were seen by male therapists than if they were all seen by female therapists. The outcomes would be the same in either case but the differential cost to the clinic would appear to be considerable. The results imply that on the basis of cost, large college counseling centers could benefit when male clients are allocated to be seen by female therapists when possible.

While outcome and length of treatment results point to fiscal rather than clinical implications, results for return rates after session one appear to be clinically relevant. Specifically, researchers often use rate of single session treatment as an index of client dropout rate. Assuming that this association with dropout is an accurate reflection of the processes active in the data set that was analyzed in this study, this translates to the observation that for every 10 female clients that drop out of therapy there are about 13 male clients. Whether the magnitude of these differences is sufficient to influence clinical or policy considerations is debatable but it would appear that efforts to reduce dropout in both male and female clients should consider whether the relatively higher rate in males (if generalizable) warrants specific consideration. While more attention to dropout for males and females appears warranted from the data, the results also indicate that the rate of single session treatment is not related to gender match between clients and therapists; i.e. matching on gender is not a solution to close the gap between males and females in this area.

A further clinical consideration relating to service utilization relates to the finding that female clients had a higher level of distress when presenting for treatment than male clients (6.58 points higher on the OQ-45 for sample 3). This was treated as a covariance/nuisance variable in our study so that analyses comparing client-therapist pairings had a control for initial client severity. However, it was surprising that the gender-stereotypical assumption that males distress threshold for seeking treatment is higher than females was not found in the sample under consideration; in fact the opposite was the case.

To summarize the primary findings of this study, female clients present to therapy relatively more distressed than males, male clients tend to spend longer in therapy with male therapists than with female therapists for similar outcomes, and the rate of single session treatment is higher in males than females. Other differences between groups on outcome or therapy duration were sufficiently small to be either statistically insignificant or clinically negligible. This study is first large sample study that we are aware of that identifies the presence of these differences but was not designed to answer the question of why they exist. If these findings prove to be robust through replication, clinicians and researchers will be faced with a number of questions to explore. For example, why are female clients typically more distressed at intake than their male counterparts, why do males tend to have a longer duration of treatment with male therapist for the same outcomes as with female therapists, and why is the rate of female clients single session treatment lower than that observed for male clients? Answers to these questions would be useful in guiding administrators and clinicians to provide services that are more accessible and efficient.

There are a number of weaknesses associated with the design of the current study that temper the strength of conclusions that can be drawn from it. Firstly, this was an uncontrolled, naturalistic study – participants were not formally randomized to treatment conditions and a range of unidentified influences could have impacted the composition...
of the comparative groups. Secondly, the study participants were drawn from a single clinical site which naturally limits the generalizability of the results without replication. Thirdly, this study was designed as an exploratory study to achieve a nomothetic appreciation of patterns within treated clients. It is quite possible that there may be ideographic trends that are obscured by this process. For example, there may be a small cluster of individual therapists who are much more effective at treating and retaining female clients as opposed to male clients, or there may be a certain diagnosis or client experience where gender matching might strongly impact outcome, or perhaps client-therapist preference matching (i.e., matching based on the preferences of the client; which was not addressed in this study) is more predictive of outcome than gender matching. The aggregating design of this study would obscure such relationships if they are present.

Although the weaknesses of this study indicate caution in interpretation, the strengths of this study address the concerns about literature in this field raised by Zlotnick et al. [11] who called for studies that employed large sample sizes, reliable outcome measures, and equivalent distribution of clients in terms of symptom severity. In all these areas the present study was able to meet these recommendations or statistically control for them (in terms of initial symptom severity).

There continues to be diverse opinions in the therapeutic field regarding the relevance of gender matching on therapy variables [6]. We would encourage clinicians to take a sophisticated and empirically grounded approach to this area as clinical assumptions are sometimes unsupported by empirical data. Research also indicates that the gender matching question is a complex one that should not be reduced to broad assumptions based on personal ideologies. Research suggests that there is wide variability both between and within subgroups of the population that defies simple conclusions regarding the utility of gender matching. The assumption of some therapists that gender matching is broadly influential on therapy outcome and the most preferred form of treatment by clients does not appear to have strong support in the research literature or the findings of this study.

If the current findings are replicated in other university based centers, then the current diversity of opinion regarding the importance of gender matching in these settings will have empirical literature as well as clinical impressions as fodder for discussion. The present study would appear to make a contribution to the available literature that provides guidance around the issue of gender matching in therapy. However, it remains to be seen whether these findings are unique to the single setting that the sample was drawn from or applicable more widely, both geographically and in the nature of clientele and therapy center.

### Table 7: Pair-wise proportion comparisons comparing non-return rates between groupings based on matching between client and therapist gender.

| Label         | Group, | $p_i$ | $n_i$ | Label         | Group, | $p_j$ | $n_j$ | $|p_i-p_j|$ | $\alpha=0.05$ | $\alpha=0.01$ |
|---------------|--------|-------|------|---------------|--------|-------|------|----------|-------------|--------------|
| Client-Therapist | Female-Female | 0.2189 | 4386 | Client-Therapist | Female-Male | 0.2096 | 5984 | 0.0093 | 0.0373 | 0.0441 |
| Female-Female  | 0.2189 | 4386 | Male-Female | 0.2852 | 2153 | 0.0663** | 0.0511 | 0.0605 |
| Female-Female  | 0.2189 | 4386 | Male-Male | 0.2678 | 4817 | 0.0499** | 0.0386 | 0.0457 |
| Female-Male    | 0.2189 | 4386 | Male-Female | 0.2852 | 2153 | 0.0756** | 0.0497 | 0.0587 |
| Female-Male    | 0.2066 | 5984 | Male-Male | 0.2678 | 4817 | 0.0582** | 0.0367 | 0.0434 |
| Male-Male      | 0.2066 | 5984 | Male-Male | 0.2678 | 4817 | 0.0174 | 0.0507 | 0.0599 |

$p_i$ = proportion for group; $n_i$ = number of cases in group

**$p<0.01$