The Association of Psychotherapist Cultural Humility and Client Experiences and Outcomes in Psychotherapy: A Meta-Analysis

Lisa Michelle Scott
The Association of Psychotherapist Cultural Humility and Client Experiences

and Outcomes in Psychotherapy: A Meta-Analysis

Lisa Michelle Scott

A dissertation submitted to the faculty of
Brigham Young University
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

Timothy B. Smith, Chair
Derek Griner
Melissa Jones
Kristina Hansen

Department of Counseling Psychology and Special Education
Brigham Young University

Copyright © 2022 Lisa Michelle Scott
All Rights Reserved
ABSTRACT

The Association of Psychotherapist Cultural Humility and Client Experiences and Outcomes in Psychotherapy: A Meta-Analysis

Lisa Michelle Scott
Department of Counseling Psychology and Special Education, BYU
Doctor of Philosophy

Psychotherapist multicultural competence has been given increasing attention in the field of psychology since the 1970s. However, individuals and communities who are Black, Indigenous, or People of Color are still less likely to attend psychotherapy and often drop out earlier than White clients. These trends highlight the importance of the ongoing research on therapist multicultural competence. There are many multicultural competencies, however, the current study focuses on therapist cultural humility. Specifically, we systematically searched all the current research on the association between cultural humility and client experiences and outcomes in therapy. Our literature search yielded eight studies that met our search criteria. Next, we synthesized the data by conducting a meta-analysis using the metric of Pearson’s $r$ as the effect size. In addition, we aggregated average item-level therapist scores and the reliability coefficient of the Cultural Humility Scale (CHS; Hook et al., 2013). We found a moderate positive correlation between client perceptions of therapist cultural humility and client experiences and outcomes in therapy of $r = .39$ ($p < .05$). Furthermore, we found the average item-level score to be 3.86 out of a maximum score of 5 (SE = .086), which corresponds with the rating of “mildly agree,” indicating that clients typically perceive therapists to demonstrate cultural humility. Finally, we found CHS to be reliable across studies, with an average alpha coefficient of .91. These finding suggest that cultural humility is an important aspect of the psychotherapeutic relationship. Furthermore, they suggest that the CHS is a reliable measure and should continue to be used in future studies. Finally, because our review yielded only eight studies, more research on this relationship is needed.

Keywords: cultural humility, Cultural Humility Scale, BIPOC, mental health professionals, psychotherapy, therapy outcomes, therapy process, therapy experiences
ACKNOWLEDGEMENTS

I believe that scholarship within psychology matters inasmuch as it shapes who we are as human beings. This understanding has guided my research questions and pursuits as well as my mentor selection throughout my time as an undergraduate and graduate student in psychology. During the course of this study the world has seen a global pandemic that took the life of my grandfather and many others, the beginning of a war in Ukraine, and within the United States, political polarization and continued violent acts of oppression and terrorism against BIPOC individuals. Within the context of these current events, I have at times felt overwhelmed by the task of becoming a thoughtful and culturally humble psychologist, researcher, and clinician. However, my mentors and friends have supported me through the challenges and detours of this process. Furthermore, my dissertation topic has granted me the opportunity to further the understanding of human nature and healing. I am deeply grateful for the opportunity to contribute to the scholarly pursuit of the understanding of cultural humility within psychotherapy, but I did not do it alone.

I first want to thank my dissertation committee chair, Dr. Timothy Smith. He has tirelessly and patiently guided me through the tedious process of conducting a meta-analysis. He shares with me a passion for asking questions and finding answers, and because of this, it has been a joy to learn from him. He is also endlessly ethical and loving and I am so grateful to have had his leadership during this time. I also want to thank my dissertation committee members Dr. Kristina Hansen and Dr. Derek Griner. Each have offered both academic and emotional support throughout my graduate experience. Their offices have been safe spaces of professional and personal development and their mentorship has been priceless to me. My BYU experience would not have been the same without them. Additionally, I would like to thank Dr. Melissa Jones, her
research pursuits have granted me many opportunities throughout my time as a graduate student, and my CV would not have as many lines if it weren’t for her guidance and support.

I also want to thank my mother, who was the first to teach me that emotions are valid and healing is always possible. She has loyally and lovingly supported me throughout my life and during graduate school. She has also edited more of my papers than all my mentors combined. She was my original mentor, and no matter how divergent our paths, continues to be my ultimate mentor in life.

Additionally, I want to extend my gratitude to my undergraduate mentor and professor, Dr. Brent Slife. Dr. Slife fearlessly forged an academic path for himself that embodied his values and faith, and because of his example, I have sought to do the same. Additionally, he opened the door of critical thinking for me which has led me to more beautiful spaces and understandings than I can list. He will stay in my heart throughout my career.

I also want to thank Daniel Crosby and his family for awarding me the Outstanding Work in Diversity Scholarship. This gift acted as a reminder that my scholarship and clinical work matter, and because of the impact this gift had on my student loans, will allow me more energy and time to focus on marginalized populations as I begin the early days of my clinical work post-graduation.

Finally, I want to thank Brigham Young University. In August I will become a doctor, not through my shear personal genius, but largely because of the community of faith that I was raised in, and their continued value for higher education. The pursuit of truth was one of the first values this community modeled for me. My life and career are what it is because of this institution and the affiliated community.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>TITLE PAGE</td>
<td>i</td>
</tr>
<tr>
<td>ABSTRACT</td>
<td>ii</td>
</tr>
<tr>
<td>ACKNOWLEDGEMENTS</td>
<td>iii</td>
</tr>
<tr>
<td>TABLE OF CONTENT</td>
<td>v</td>
</tr>
<tr>
<td>LIST OF FIGURES</td>
<td>ix</td>
</tr>
<tr>
<td>DESCRIPTION OF DISSERTATION STRUCTURE AND CONTENT</td>
<td>x</td>
</tr>
<tr>
<td>Introduction</td>
<td>1</td>
</tr>
<tr>
<td>Statement of the Problem</td>
<td>2</td>
</tr>
<tr>
<td>Statement of the Purpose</td>
<td>3</td>
</tr>
<tr>
<td>Research Questions</td>
<td>3</td>
</tr>
<tr>
<td>Method</td>
<td>4</td>
</tr>
<tr>
<td>Literature Search</td>
<td>4</td>
</tr>
<tr>
<td>Inclusion Criteria</td>
<td>4</td>
</tr>
<tr>
<td>Data Coding</td>
<td>5</td>
</tr>
<tr>
<td>Computation of Effect Size Estimates</td>
<td>5</td>
</tr>
<tr>
<td>Data Analyses</td>
<td>6</td>
</tr>
<tr>
<td>Moderation by Psychotherapist and Client Variables</td>
<td>6</td>
</tr>
<tr>
<td>Publication Bias</td>
<td>7</td>
</tr>
<tr>
<td>Results</td>
<td>7</td>
</tr>
<tr>
<td>Search Results</td>
<td>7</td>
</tr>
<tr>
<td>Interrater Reliability</td>
<td>10</td>
</tr>
<tr>
<td>Descriptive Statistics</td>
<td>10</td>
</tr>
</tbody>
</table>
Psychological Explanations for Mental Health Professionals’ Cultural Biases and Lack of Cultural Humility

Internal Psychological Processes

Implicit Bias

Bias Blind Spot

Ingroup Bias

Empathy Gap

Outgroup Homogeneity Effect

Fundamental Attribution Error

Ultimate Attribution Error

Just World Phenomenon

Continued Influence Effect

Mere Exposure Effect

Social and Group Dynamics

Social Identity Theory

Realistic Group Conflict Theory

Optimal Distinctiveness Theory

Aversive Racism

Contact Hypothesis

Psychotherapist Multicultural Competence: A Solution to Innate Cognitive and Social Cultural Biases

Multicultural Competencies in Psychotherapy
LIST OF FIGURES

Figure 1 Search Process and Findings .................................................................................................................. 8
Figure 2 Individual Study Effect Sizes ................................................................................................................. 12
Figure 3 Funnel Plot ........................................................................................................................................... 14
DESCRIPTION OF DISSERTATION STRUCTURE AND CONTENT

This dissertation, *The Association of Psychotherapist Cultural Humility and Client Experiences and Outcomes in Psychotherapy: A Meta-Analysis*, is formatted according to the MSE Journal-Ready style. The preliminary pages align with BYU’s requirements for submission. Once the authors choose a journal to submit the article to, it will be more specifically formatted according to the requirements of that journal.

The appendices include the literature review (Appendix A), Cultural Humility Scale (Appendix B), and data coding sheet (Appendix C). Appendix A includes the full-length literature review completed prior to the completion of the study. Appendix B includes the Cultural Humility Scale. This meta-analysis reviewed the association of this scale with various therapy experiences and outcomes of BIPOC clients, and we therefore included the full measure for reference. Finally, Appendix C includes the data coding sheet used by the research team to code the studies that were retrieved during our systematic search of the literature. This coded data was then statistically analyzed.

This dissertation includes two references lists. The first reference list includes all references included in this dissertation (both the main article and those found in the literature review). The second reference list is located after the literature review in Appendix A and includes only the references used in the review.
Introduction

Many racial disparities persist in the United States. Individuals and groups who are Black, Indigenous, or People of Color (BIPOC) have experienced systematic disadvantages across time and across institutions, including social and health services (Feagin, 2013; Feagin & Bennefield, 2014; Kendi, 2017). Unfortunately, racial disparities have also characterized psychology (Atkinson et al., 1979; Guthrie, 1976; Samuda, 1998).

In the United States, BIPOC are less likely than Whites to utilize mental health services (Atkinson et al., 1979; Gallo et al., 1995; Smith & Trimble, 2016), even when the rates of mental illness are similar across groups (Regier et al., 1993). For instance, one study with a nationally representative sample found that a majority of Latinx, Asians, and African Americans who had experienced depression in the last year had not accessed any mental health treatment while the majority of Whites with the same condition had (Alegría et al., 2008). Even when BIPOC seek treatment, some research has shown that the quality of care is lacking in contrast to the quality received by White individuals (Alegría et al., 2008; Blanco, 2010; Carson et al., 2014). In addition, retention rates are also lower for these populations (Allen et al., 2016; Fortuna et al., 2010; Sue et al., 1994).

There are many factors at play in the disparities of access and quality treatment for BIPOC. For instance, BIPOC at times deal with culture-based mental health stigma (Park et al., 2018; Pedersen & Paves, 2014; Samuel, 2015; Turner et al., 2015). In addition, BIPOC are more likely than Whites to experience poverty, and therefore may experience barriers in paying for treatment (U.S. Census Bureau, 2019). While there may be a variety of barriers contributing to the differences in utilization rates across races, the field must address how psychotherapist actions and attitudes might be involved in these disparities. If so many BIPOC forego treatment
in spite of a need. What is playing out in the therapy room that they might be avoiding? Is there something that psychotherapists can do to make therapy more accessible and appealing to these populations?

Humans are prone to believe that their own perceptions are more accurate than the perceptions of others (Pronin et al., 2002). That dynamic has implications for racial disparities in psychology, with psychotherapists and other mental health professionals (MHPs) possibly believing that their own expert opinions are more accurate than the experiences reported by clients historically underrepresented and often misunderstood by mental health professionals (Sue et al., 1982). Psychologists increasingly recognize this dynamic, and the field as a whole has taken systematic steps to correct blatant racial inequities (Arredondo & Perez, 2006; Smith & Trimble, 2016; Sue et al., 1982). Although psychologists today receive training in multicultural competence (Smith et al., 2006), it is important to continue to investigate the ways MHPs may continue to overlook or dismiss clients’ cultural experiences, even unintentionally or even after being aware of the need to counter such biases.

Statement of the Problem

Recently, scholars have defined psychotherapist cultural humility as a multicultural counseling competence that involves proactively seeking out clients’ cultural experiences and aligning therapy with those experiences (in contrast to the psychotherapist taking the lead in session, moving forward without ascertaining clients’ cultural values, etc.; Hook et al., 2013). Thus far, studies of psychotherapist cultural humility have demonstrated strong positive correlations with BIPOC client experiences and outcomes in therapy (Hook et al., 2016; Owen et al., 2016), but that work has not yet been synthesized. Dozens of reviews and a few meta-analyses have addressed the broader issue of psychotherapist multicultural competence (Soto et
al., 2018; Tao et al., 2015; Worthington et al., 2007), but to our knowledge none has specifically addressed the degree to which clients experience different psychotherapist evaluations and treatment outcomes based on the cultural humility of their psychotherapist.

**Statement of the Purpose**

This proposed meta-analysis seeks to synthesize all research that has investigated the association between perceived psychotherapist cultural humility and BIPOC clients’ experiences and outcomes in treatment. We will also analyze the reliability of the Cultural Humility Scale or CHS (Hook et al., 2013) across studies. We hope this data will raise professional awareness about what may be needed to move forward toward more multiculturally competent treatment, as well as provide data about the utility of the CHS as a research measure.

**Research Questions**

This meta-analysis will evaluate the following:

1. The weighted average score on the CHS, which indicates the degree to which clients in research studies perceived cultural humility demonstrated by their psychotherapists.

2. The weighted average reliability coefficient of the CHS across studies, which indicates the degree to which participants in research studies responded consistently to CHS items.

3. The magnitude of the association between CHS scores and clients’ experiences and outcomes in treatment.
Method

Literature Search

The literature search consisted of four methods. First, researchers searched the database Google Scholar for all articles citing the CHS. Second, researchers searched the reference lists of retrieved articles. Third, researchers emailed authors and coauthors who published one or more of the gathered articles to inquire if they had presented on or published any additional research on the topic. Finally, researchers searched for studies difficult to locate, known as the grey or fugitive literature (Rothstein & Hopewell, 2009), such as unpublished studies (e.g., conference presentations). Researchers did this by using the database PsycEXTRA and by searching through conference programs of the American Psychological Association (APA), American Counseling Association (ACA), and National Association of Social Workers (NASW) for the years 2015-2019.

Inclusion Criteria

To be included in the meta-analysis, studies must have reported a quantitative association of the CHS (Hook et al., 2013) with a psychotherapy process or outcome measure administered to BIPOC clients in the U.S. or Canada. The geographic location was limited because of the differences in cross-cultural relations from country to country. Studies were also restricted to mental health treatments, not including school counseling, career/vocational counseling or substance abuse counseling. This is because of the differences in methodology across these fields. Because the CHS was created in 2013, studies were restricted to the years 2014-2020. Studies were included if they were written in any of the languages spoken by the research team, which includes English, Spanish, Portuguese, and French. Searching continued until June 2020.
Data Coding

Coding was conducted by two teams of two researchers each, with every article coded independently by both teams. Inconsistencies between coding teams were addressed and resolved through further review of the manuscript by one person from each of the original coding teams. Furthermore, if these coders could not come to a consensus on a variable, the discrepancy was brought to the team meeting for the team to adjudicate. Variables that were coded included basic study information such as whether the study was published and in which field it was published, demographics and basic information about the psychotherapists and clients involved, whether the psychotherapists had multicultural training, treatment type, information about the study design, whether the effect size was based on the whole sample or subgroups, statistics used, dependent variable and its reliability, internal consistency coefficient of the independent variable, the effect size and sample size. Finally, we measured interrater agreement after all coding had been completed.

Computation of Effect Size Estimates

The included studies used a variety of statistics to measure the association between CHS and the other variables. These statistics included Pearson’s $r$, beta weights, path coefficients and $p$ values. To compare data across studies, all statistics were converted to Pearson’s $r$. Some of the studies reported more than one dependent measure, such as including both an outcome measure and a therapeutic alliance measure. To ensure independence of the sample data (Cooper, 1998; Cooper & Hedges, 1994; Hedges & Olkin, 1985), and to ensure that studies with multiple dependent measures did not have a greater impact on the final statistic, the effect sizes of these studies were averaged (weighted by number of participants in each effect size) to create an
aggregate effect size for the study. This was done to ensure that each study population only contributed one effect size to the omnibus analysis.

**Data Analyses**

In this meta-analysis, the studies that were aggregated involved different samples, methods, measures, etc., such that the results varied across studies. Given these differences, analyses were conducted using inverse variance weighted random effects models. This procedure is in line with research recommendations (Field, 2005). Random effects models do not assume that data are consistent across studies but rather that multiple factors influence study findings. Random effects models allow the results of the meta-analysis to be generalized to future research.

Correlation coefficients were transformed to Fisher's Z for purposes of aggregation and then transformed back to the metric of Pearson's r for purposes of interpretation. Following the calculation of the overall effect size, effect size heterogeneity was examined by first calculating a Q statistic and then an $I^2$ statistic. Statistical significance was set at $p$ is less than or equal to .05.

**Moderation by Psychotherapist and Client Variables**

Participant characteristics and study characteristics were considered as potential moderators of the association between cultural humility and client experiences in therapy. Different analyses were employed for the continuous and categorical variables. For continuous data, random effects weighted simple regression models were run between the variable of interest and the effect size obtained from the study. Categorical data was analyzed using random effects weight analyses of variance (ANOVAs).
Publication Bias

Meta-analyses risk reporting inflated effect sizes because of what is called publication bias. This bias often occurs because meta-analyses typically include more published than unpublished data, and because published data often include larger effect sizes than unpublished data. In order to account for this bias, we (a) calculated a fail-safe N (Begg, 1994), (b) analyzed a scatter-plot displaying effect sizes (x-axis) and the number of participants by study (y-axis), and (c) used the “trim and fill” method by Duval and Tweedie (2000a, 2000b).

A fail-safe N is the theoretical number of unpublished and/or missed studies whose effect sizes would equal zero and therefore cause the omnibus effect size to reduce to a trivial number according to Cohen’s (1988) guidelines. After the fail-safe N was calculated, we examined a scatter-plot displaying effect sizes (x-axis) and the number of participants by study (y-axis).

Finally, the “trim and fill” method was used to assess the number of missing studies. In this method, asymmetrical studies are trimmed (i.e., those that lie on the outer edges of the scatter-plot that did not have corresponding data points on the opposite side of the plot) and then filled with equivalent values to obtain a new mean effect size. This process was repeated until the distribution was symmetrical, or in other words, the data points become evenly scattered around the mean. After the distribution was deemed symmetrical, the resulting estimate was reported as a possible correction to likely publication bias.

Results

Search Results

Researchers searched for articles that fit our inclusion criteria. The first database searched was Google Scholar. A search of the years 2014 through 2020 yielded a total of 548 articles that cited the CHS (Hook et al., 2013). After the research team filtered these articles through the
inclusion criteria, a total of 12 articles were downloaded in pdf form and added to our database for further inspection. This final inspection included the research team closely reviewing the full-length articles and determining together if they should be included or excluded. The second database searched was APA PsycEXTRA. A search of the years 2014 through 2020 yielded a total of 274 studies which mentioned the concept of cultural humility, with 8 articles added to the database for further inspection by the research team. See Figure 1 for a visual depiction of the entire search process.

**Figure 1**

*Search Process and Findings*
As described in the Method section, articles were excluded if they did not include any data, were qualitative, were not conducted in North America, were not correlated with psychotherapy outcome/process variables, or used the CHS to measure therapist humility related to some characteristic other than race and/or culture of the client. We sought for prior meta-analyses or systematic reviews as a source of potential studies, and we excluded identical articles identified across different sources. Once we had completed our search, the research team analyzed the gathered articles as a group and decided if any more should be excluded based on our exclusion criteria. After this process, we retained a total of nine articles. One study was in the process of re-examining data after finding errors, so it was unavailable. Authors followed up, but the data remained unavailable. Because of this, our final included number of studies was eight. Each of these eight studies contributed one data point to our analysis.

In addition to the general search, we were also interested in understanding the reliability of the CHS so that we might provide information about the utility of the CHS as a clinical research measure. Furthermore, we coded the average item-level scores because we wanted to begin to understand how clients typically experience their therapist’s cultural humility. To do this, two team members searched for any articles that fit our criteria, but instead of only including articles that included correlations between therapy process/outcomes and CHS, they included any articles that reported CHS reliability and item-level ratings. The search for articles with CHS reliability produced three more articles in addition to the eight already included in our review, which therefore provided a total of 11 articles to code for these variables. The search for the average item-level score yielded four more articles, which brought the total to 12 for that variable.
Interrater Reliability

After coding was completed, the interrater reliability was calculated for both the continuous and categorical variables. For the categorical variables, the average interrater reliability calculated using Cohen’s Kappa was 0.716. Cohen (1960) suggests values below or equal to zero to indicate no agreement, values between 0.01–0.20 indicating none to slight agreement, 0.21–0.40 indicating fair agreement, 0.41–0.60 indicating moderate agreement, 0.61–0.80 indicating substantial agreement, and 0.81–1.00 as almost perfect agreement. Based on these guidelines, we consider our interrater agreement to be substantial.

For the continuous variables, the average was calculated using the intraclass correlation coefficient, which was 0.873. We used the rating system suggested by Koo and Li (2016) to determine the reliability of this coefficient. These authors consider values below 0.5 to have poor reliability, values between 0.5 and 0.75 to have moderate reliability, values between 0.75 and 0.9 to have good reliability, and values greater than 0.90 to have excellent reliability. Based on this guideline, the research team considered this coefficient to mean that the inter-rater reliability was “good.” Our research design was specifically set up to produce this reliability. For instance, when coding discrepancies occurred between coding teams, these were addressed and resolved through further review of the manuscript by one person from each of the original coding teams. Furthermore, if these coders could not come to a consensus on a variable, the discrepancy was brought to the team meeting for the team to adjudicate.

Descriptive Statistics

Across the eight studies included in the meta-analysis, the total number of participants was 4,170. Four studies reported the age of the therapist, which ranged from 26 to 43, with an average age of 38 years. Six studies reported the gender of the therapist, which ranged from 64%
to 82% females, with an average of 75% females. Seven studies reported the race of the therapist, with an average of 65% White/European Americans, 11% Black/African Americans, 5% Asian Americans, 2.6% Latinx Americans, and 1% Native Americans. The percentage of other races ranged from 3 to 43%, with an average of 15%. All of the studies reported the age, gender, and race of the client. The age of the clients ranged from 23 to 35, with an average of 27 years. The gender of the clients ranged from 21% to 100% females, with the average of 63% females. The studies reported an average of 43% Black/African American clients, 17% Latinx clients, 16% Asian American clients, 6% White/European American clients, and 1% Native American clients. Other races and biracial or multiracial clients made up an average of 16% of clients.

The studies were also coded for geographic location, treatment type, study design (dependent variable), and random selection. Seven of these studies were from the field of psychology, while one was from the field of counseling. Three of the studies were national surveys. Three of the studies were conducted in the Southern United States, and one study was conducted in the Western United States. Three of the studies included only individual psychotherapy as their treatment modality, while five or 62% included various modalities. Two, or 25%, of the studies were retrospective surveys; five, or 63%, were cross-sectional surveys with current clients; and one study was longitudinal. All the studies included convenience samples. Lastly, one of the studies included a dependent variable of therapist microaggressions; six, or 75%, included a dependent variable that was a therapy outcome measure; and one included a dependent measure that was the client’s perceptions of working alliance with the therapist.
**Omnibus Effect Size Correlation**

To estimate the overall association between client ratings of therapist cultural humility and client experiences/outcomes in therapy, a random effects model was used to calculate the omnibus effect size correlation. The resulting value was $r = .39 \ (p < .05)$ with a 95% confidence interval of $r = .140$ to $r = .593$. The effect size estimates ranged from -.34 to .751. The one effect size having a negative direction was extracted from a study whose sample included only jail inmates. Overall, the effect size estimates were characterized by a very large degree of heterogeneity ($I^2 = 96.5$, $\tau^2 = 0.1$, $Q = 105.1$, $p < .05$). The data are displayed in a forest plot in Figure 2. Given the heterogeneity of the findings, the researchers therefore conducted analyses of potential moderating variables.

**Figure 2**

*Individual Study Effect Sizes*

<table>
<thead>
<tr>
<th>Study</th>
<th>Correlation with 95% CI</th>
<th>Weight (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Davis 2016</td>
<td>0.62 [ 0.50, 0.72]</td>
<td>12.63</td>
</tr>
<tr>
<td>DeBlare 2019</td>
<td>0.31 [ 0.20, 0.41]</td>
<td>13.04</td>
</tr>
<tr>
<td>Hook 2013</td>
<td>0.57 [ 0.44, 0.68]</td>
<td>12.61</td>
</tr>
<tr>
<td>Hook 2016</td>
<td>0.26 [ 0.22, 0.30]</td>
<td>13.31</td>
</tr>
<tr>
<td>Owen 2014</td>
<td>0.31 [ 0.22, 0.50]</td>
<td>11.45</td>
</tr>
<tr>
<td>Owen 2016</td>
<td>0.39 [ 0.23, 0.53]</td>
<td>12.61</td>
</tr>
<tr>
<td>Owen 2019</td>
<td>-0.34 [ -0.51, -0.15]</td>
<td>12.41</td>
</tr>
<tr>
<td>Wright 2019</td>
<td>0.75 [ 0.62, 0.84]</td>
<td>11.94</td>
</tr>
</tbody>
</table>

**Overall**

Heterogeneity: $\tau^2 = 0.14$, $I^2 = 96.49\%$, $H^2 = 28.47$

Test of $\tau = 0$: $Q(7) = 105.14$, $p = 0.00$

Test of $\theta = 0$: $z = 2.98$, $p = 0.00$

Random-effects REML model
**Moderator Analyses**

To determine whether certain variables impacted the effect size of the studies, we ran moderation analyses separately for continuous and categorical variables coded in this study. Continuous variables included the age, education level, percentage of BIPOC clients, percentage of females, and SES of the therapist and the client, as well as the total sample size, reliability coefficient of the CHS in that study, and study mean CHS score. To analyze all continuous variables which had been coded, we ran separate random effects weighted meta-regressions for each variable, one at a time. None of the results reached statistical significance.

We next analyzed categorical variables that had been coded including type of sample of clinicians (students, psychologists, etc.), type of sample of clients, location of the study, treatment type, random selection of participants, study design, dependent variable type, and who provided dependent variable data. To analyze these variables, we ran separate random effects weighted ANOVA equivalent for each variable. None of the results reached statistical significance.

**Average Cultural Humility Scale Reliability**

Researchers found 11 studies that reported the reliability coefficient for the CHS. For this analysis, it did not matter if the study included the correlation of the CHS and measures of therapy process or outcomes. Therefore, three additional studies were included in this aspect of the analysis. The average CHS reliability coefficient was calculated using an average random effects model and shows the measure to be quite reliable with a coefficient of $k = .91$ ($p < .0001$). The 95% confidence interval was .90 to .92. Next, researchers calculated the average for CHS item-level ratings.
Average Cultural Humility Scale Item-Level Ratings

The researchers found 12 studies that included the mean and standard deviation of CHS item-level ratings. The average score for the items was 3.86 out of a maximum score of 5 (SE = .086), which corresponds with the qualitative rating of “mildly agree.” The 95% confidence interval was 3.83 to 4.09. The average standard deviation across studies was 0.71.

Publication Bias

To determine whether our meta-analysis was impacted by publication bias, we ran an Egger’s Regression, Begg’s Test, analyzed the funnel plot, and employed the “trim and fill” method. The funnel plot showed one or two studies that were imbalanced in the distribution, possibly indicative of publication bias. The funnel plot can be viewed in Figure 3. Egger’s regression and Begg’s test both produced results that were not significant ($p > .05$), indicating that publication bias was unlikely. Finally, we employed the “trim and fill” method, and no studies were imputed. Based on these analyses, we concluded that publication bias was an unlikely threat to the result of this meta-analysis.

Figure 3

Funnel Plot
Discussion

Prior research had concluded that BIPOC clients encounter more barriers in mental health treatment than do White individuals (Alegría et al., 2008; Pedersen & Paves, 2014; Smith & Trimble, 2016), including therapists’ racial and cultural bias in diagnosis (Becker et al., 2003; Cranford, 1999; Kaden, 2010), treatment recommendations (Cohen et al., 2003; Gordon et al., 2006; Kugelmass, 2018), and clinical reactions (Blanco, 2010; Rieffel, 2006). These biases are evidence of the need for MHPs to display cultural humility as they work with BIPOC clients. Because of these realities, the important work of understanding this specific aspect of multicultural competence must continue, and we feel that our meta-analysis has provided important data to support that work.

Our analyses revealed a moderate positive correlation between client ratings of their therapist on the CHS and client experience and outcome variables, suggesting that professionals and training programs can address therapist cultural humility as an important consideration when working with BIPOC clients. In addition, we found that clients tended to report that their therapists displayed a mild degree of cultural humility. And finally, we found that the CHS was a reliable measure across studies. These results provide valuable contributions to the current research on multicultural competence in therapy.

While our findings provide more confidence in the importance of cultural humility in therapy, potentially the most striking finding of this analysis was the lack of research on the CHS to date. Our search yielded 800 studies that cite the CHS, and yet, only 8 of these inspected the relationship between the CHS and therapy outcomes. Furthermore, of these eight, only six are peer-reviewed and published. The topic of cultural humility in therapy is a popular topic within
the field (800 articles popular), and therefore should be taken seriously as researchers look to their future agendas.

Furthermore, given that clinicians are themselves humans capable of oppression, it should be a priority for the field to conduct research that will highlight the best ways to address this reality in the clinician-client relationship. BIPOC individuals have the right to receive mental health support in environments that will not inflict more harm. It is our belief that more research into the concept of therapist cultural humility is an important aspect of ensuring that this right is protected.

**Overview of Main Effects**

We identified eight studies reporting the association between client experiences in psychotherapy and the clients’ ratings of their therapists’ cultural humility on the CHS. On average, those studies reported a moderately positive correlation of $r = .39$ ($p < .05$) with a 95% confidence interval of $r = .140$ to $r = .593$. In other words, when clients perceive their therapist as demonstrating cultural humility, they are also more likely to report having more positive experiences in psychotherapy (and vice versa). This correlation aligns nicely with previous meta-analyses on the relationship between client-perceived multicultural competence of the therapist and therapy outcome which included omnibus correlations of $r = .38$ (Soto et al., 2018) and $r = .29$ (Tao et al., 2015).

Of the eight studies included in our meta-analysis, all but one showed a positive correlation between the two variables. Owen et al. (under review) reported a -.34 correlation between therapy outcomes and cultural humility. However, this study evaluated a very specific population: jail inmates. It is plausible that this population perceives therapist attributes in a different way than clients not coerced by their circumstances to engage in psychotherapy.
On the other end of the distribution, one study (Wright, 2019) reported a very strong correlation of $r = .66$. When correlations of such large magnitude occur, it raises questions about whether clients sufficiently distinguished between therapist cultural humility and their overall interactions with the therapist (in this case, measured by the Working Alliance Inventory; Falkenström et al., 2015). Nevertheless, it is reasonable that perceptions of therapist cultural humility would be interrelated with perceptions of alliance with a therapist. Overall, although the data were correlational and not causal, the observed associations provide preliminary support for a recommendation for therapists to demonstrate cultural humility in their work.

**Overview of Average Cultural Humility Scale Item-Level Ratings**

To understand how clients typically experience their therapists in regard to cultural humility, we calculated the average item-level score per participant (with reverse-scoring on certain items taken into account). This was found to be 3.86 out of 5 possible points, corresponding with a rating of “Mildly Agree.” This indicates that, on average, clients experienced their therapists as demonstrating some cultural humility. This average was based on a total of 12 studies and 4,219 participants. The fact that study means ranged from 3.45 to 4.61 further instills confidence in the interpretation that therapists were generally rated somewhat positively, although variability clearly exists (average $SD = 0.7$).

This finding stands in contrast to the many studies which have shown racial and/or cultural bias in mental health treatment (Becker et al., 2003; Blanco, 2010; Cohen et al., 2003; Cranford, 1999; Gordon et al., 2006; Kaden, 2010; Kugelmass, 2018; Rieffel, 2006). It will likely be important for future research to investigate the relationship between perceived cultural humility and unbiased treatment.
Overview of Cultural Humility Scale Reliability

We were also interested in providing information about the reliability of the CHS as a clinical research measure. To do this, we calculated the average reliability coefficient of the CHS across studies. The average random effects weighted internal consistency reliability coefficient of the CHS was found to be .91, which indicates that the CHS scores tended to be adequately reliable across studies. This average was based on data from 11 studies and 4,125 participants. Given the observed results, it appears that researchers can be confident in using the CHS scale in their future research.

Summary of Moderator Variables

To evaluate whether any environmental, client or therapist variables impacted the correlation between the two measures, we coded and analyzed potential moderator variables. Our analysis did not reveal any statistically significant moderator variables. Therefore, at this time, we cannot claim to understand any variables that might impact the relationship between CHS and therapy experiences and outcomes. This is likely due to the small sample of studies that we were able to obtain in our search, which resulted in very low statistical power.

Strengths of the Meta-Analytic Methodology Used

This study used meta-analysis to better understand the relationship between CHS and client experience and outcome variables. Because we employed a meta-analytic technique, we were able to obtain information, such as an omnibus effect size, that a typical literature review could not have obtained. In addition, this effect size is based on a much larger sample than any of the individual studies, therefore inspiring more confidence in its accuracy. Furthermore, because we have included studies with different variables measuring experience and outcome of therapy, the overall topic is broader than an individual study could examine. For example, while Wright
(2019) studied the relationship between “working alliance” and the CHS, we can include “working alliance” as just one of the many variables under the larger construct of “client experiences and outcomes of therapy.” Finally, because of the exhaustive and precise nature of our literature search, our analysis can confidently show the current state of the research related to our initial research questions.

**Limitations of the Meta-Analytic Methodology**

While meta-analysis has many strengths, it also has certain limitations. For example, the only type of data that can be analyzed with a meta-analysis is quantitative data. Therefore, case studies, theoretical pieces, and qualitative research were not analyzed or synthesized into our findings. There may be important findings and insights that could be gained from these types of research projects, but they are outside of the scope of a meta-analysis. For those wishing to gain a broader understanding of cultural humility as a construct, these sources may be valuable.

In addition, because meta-analytic findings are based on the research work of other research teams, the quality of our findings is dependent on the quality of the studies we included. Furthermore, within our sample of studies, two out of the eight were unpublished. Because of this, these studies have yet to be peer reviewed and therefore receive systematic quality checks. Another limitation is the small sample of studies included in our meta-analysis. To reach a more robust conclusion about the relationship between CHS and therapy experiences and outcomes, more research needs to be done, and a future meta-analysis would be beneficial.

Another limitation of our data is the voluntary nature of our self-report measure. Because of this, the sample of participants may have been restricted. This is because only clients who were willing to take the questionnaire and respond to questions related to their therapist’s cultural humility were included.
Finally, our meta-analysis examined correlation coefficients. Because of this, we cannot make any interpretations about cause and effect. In other words, our findings do not allow us to conclude that perceived cultural humility of the therapist causes quality therapy experiences and outcomes, or vice versa. Our analysis is a useful start in understanding this relationship, but more research should be done.

**Implications for Clinical Practice**

Our findings showed that clients, on average, offer their therapists high scores on the CHS. In other words, most clients in our sample experienced their therapist as demonstrating acceptable levels of cultural humility. This is good news for psychotherapists, accrediting bodies, and training programs. MHPs are likely engaging with the construct of cultural humility in important ways. However, given the nature of the construct and the reality of oppression for BIPOC clients, clinicians should approach the topic with continual curiosity about the ways in which they can improve in their humility.

Furthermore, given the various reasons individuals of minoritized racial, ethnic or cultural identities might distrust the process of psychotherapy (Becker et al., 2003; Blanco, 2010; Cohen et al., 2003; Cranford, 1999; Gordon et al., 2006; Kaden, 2010; Kugelmass, 2018; Rieffel, 2006), this data may be useful in the psychotherapeutic outreach and recruitment efforts for individuals in these groups. If individuals in need of psychotherapy services are educated on the probability of working with a therapist who is displaying cultural humility, this may increase their chances of seeking out and participating in these vital services.

Our finding of a moderately positive correlation between psychotherapy experiences and outcomes suggests that it may be useful to expose MHPs in training to the construct of cultural humility as defined by Hook et al. (2013). In fact, our findings align with Soto and colleague’s
(2018) meta-analysis that found that client-rated measures of therapist cultural competence correlated strongly ($r = 0.38$) with therapy outcomes. It seems that awareness of cultural aspects to treatment are likely a very important aspect of clinical training. In addition, it will likely be beneficial for future research to explore the specifics of how supervising clinicians might support supervisees in building the specific competency of cultural humility.

As a starting point, the CHS can act as a framework with which to begin training supervisees. For instance, item eight on the scale asks clients to rank the statement, “Regarding the core aspect(s) of my cultural background, my counselor makes assumptions about me.” This item could be used to highlight the difference between learning about a client’s cultural background and assuming that what you learn will always apply to them. Indeed, each item on the CHS could be discussed, explored and even used as material for clinical roleplays. It would likely also be useful to include the CHS in outcome measures given to the clients of new therapists. This would allow supervisors to track clinician cultural humility and become aware when a trainee is struggling in this area.

Implications for Future Research

We consider our findings valuable, but given the small nature of our study, we encourage continued research in this area. Luckily, the overall reliability of the CHS was shown to be robust. Because of this, researchers can have confidence in the use of CHS in their research. Furthermore, the small size of our study kept us from locating any potential moderation in the data, and therefore more research highlighting potential moderator variables is needed. As more research is published in this area, it may be beneficial to conduct a future meta-analysis on the association between CHS and the psychotherapeutic process and outcomes.
Our meta-analysis looked specifically at the multicultural competency of cultural humility. However, more research should be conducted on the various multicultural competencies. Specifically, it would likely be useful for future meta-analyses to be conducted for each of these various constructs. It may also be useful for future studies to investigate the interactions of these competencies within psychotherapy.

In addition, as mentioned in our introduction, retention rates (Allen et al., 2016; Fortuna et al., 2010) and quality of care (Alegria et al., 2008; Blanco, 2010; Carson et al., 2014) in psychotherapy have been shown to be worse for BIPOC individuals. Because of this, it would likely be useful for future research to investigate the relationship between perceived cultural humility of the therapist and quality of care and retention rates for BIPOC clients.

Another potential area for research is the comparison of experiences of cultural humility for White individuals and BIPOC individuals. While our meta-analysis focused on the experiences of BIPOC individuals, it may be useful to investigate how these experiences differ across racial and cultural groups. This could further highlight which groups are being underserved in psychotherapy.

Given the positive correlation between cultural humility and therapy outcomes, it is likely important for us to learn about the barriers involved in practicing cultural humility. Another potential area for investigation is an analysis of characteristics of MHPs who struggle with cultural humility. These characteristics may be useful in supporting these MHPs in the cultivation of cultural humility.

Furthermore, our research highlights a positive correlational relationship between CHS and therapy experience and outcome, but it does not explore why this correlation exists. It is likely important for future researchers to interview clients about their positive and negative
experiences with therapists regarding cultural humility. This type of data could highlight how cultural humility impacts the therapy process and could provide first-hand accounts of cultural humility in practice.

A final recommendation would be further investigation into the reasons for Owen et al. (under review) result of a negative correlation between CHS and the psychotherapeutic measures used in the study. Qualitative studies may help to reveal the unique perspectives of jail inmates regarding therapist cultural humility. Furthermore, future studies may be useful in locating other populations that have worse experiences in therapy when their therapist displays cultural humility.

**Conclusion**

Negative impacts of racial, ethnic, and cultural biases are clear. Because of this, it is more important than ever for clinicians to be working toward a culturally humble attitude so that they can offer a vital health service to those impacted by systemic oppression. Our analysis found evidence that this attitude is related to positive experiences and outcomes for BIPOC clients, and that therapists, on average, are likely doing well in this area. However, because of the small number of studies included in this analysis, we encourage further research.

Our study also found the CHS to be a reliable measure, and we therefore recommend its continued use in future research. We hope that future research offers further insight about therapist cultural humility and the ways it is related to and impacts therapeutic experiences and outcomes for BIPOC clients.
References


Guthrie, R. V. (1976). *Even the rat was white: A historical view of psychology* (Classical Ed.). Pearson.


https://doi.org/10.1037/cdp0000047


https://www.census.gov/programs-surveys/cps.html


APPENDIX A

Literature Review

Racial and Cultural Bias in Early Psychology

During the nineteenth and early twentieth century, most subjects in psychological research were White, and when BIPOC were included, it was often to find inferior characteristics, mental abilities, and personality traits (Guthrie, 1976; Richards, 1997; Winston, 2004). In fact, nineteenth century research produced hundreds of these studies. One such research project was the study of eugenics, which ultimately impacted the lethal thinking and actions of the Nazis in World War II (Guthrie, 1976). In addition, many of these racist and ethnocentric “scientific findings” often impacted U.S. culture and policy. In fact, this was blatantly occurring as late as 1973, when Henry Garrett (a former APA president) used erroneous research to show African American intellectual inferiority to justify racial segregation (Guthrie, 1976).

In addition to research that sought to “prove” racial minority inferiority, there existed a general cultural assumption within the field that Western psychologists were gaining knowledge that was unbiased and could be universally applied across cultures (Dawson, 1971). Issues of culture, race, ethnicity, gender, religion and spirituality, and sexual orientation, were rarely considered prior to the second half of the 20th century (Smith & Trimble, 2016). In short, the early history of psychology could best be characterized as systemic cultural arrogance, the diametric opposite of cultural humility emphasized in contemporary multicultural psychology. As concerns about multicultural competence began to be raised (mostly by scholars of color), scholars in the social sciences began to produce research that considered cultural factors and contradicted the biased findings based on the presumption of cultural inferiority (Guthrie, 1976;
Sue et al., 1982). For instance, Horas Man Bond at Langston University noticed that many of the studies that found Black individuals to be intellectually inferior to White individuals had sampled individuals only from the labor class (Guthrie, 1976). Because of this, he gathered Black children from middle and professional homes and tested their intelligence. He also worked diligently to create good rapport with his subjects. Sixty-three percent of his subjects received scores above 106—successfully contradicting the previous findings of Black intellectual inferiority. His research was one project among many that began to push back against the biased findings produced from culturally arrogant conceptions of ethnic and BIPOC (Guthrie, 1976). These studies helped to pave the way for the multicultural movement that would eventually surface within the field.

The Multicultural Movement in Psychology

The Civil Rights movement had a profound influence on the field of psychology and marked the rise of multicultural psychology and counseling (Arredondo & Perez, 2006; Smith & Trimble, 2016; Sue et al., 1999). The 1970s through the 1990s was a period of tremendous growth for the field as multicultural journals, conferences and organizations began to form (Arredondo & Perez, 2006; Smith & Trimble, 2016; Sue et al., 1999). Part of this change was also due to the rise in BIPOC and women receiving graduate degrees and entering the field (Smith & Trimble, 2016). These changes enriched the field with more diverse perspectives and research endeavors.

A landmark event during this time was the formation of the Association of Black Psychologists (Arredondo & Perez, 2006). The creation of this association influenced the formation of the Asian American Psychological Association, the National Hispanic Psychological Association (or the National Latina/o Psychological Association), and the Society
of Indian Psychologists in the 1970s and early 80s (Arredondo & Perez, 2006). In addition, scholars during this time began to publish articles and books outlining the ethnocentrism within the field (Arredondo & Perez, 2006). Furthermore, in 1982, the first official multicultural guidelines document was written by Derald Wing Sue entitled “Position Paper: Cross-Cultural Counseling Competencies” (Sue et al., 1982). The publication of this document ushered in a series of other documents outlining multicultural competencies.

Starting in the late 1980s, APA required every accredited doctoral program to include multicultural competencies in doctoral coursework (Korman, 1974). Eventually, the APA published specific guidelines for multicultural practice and training in 2002 that were updated in 2017 (APA, 2017). As exemplified by these systematic changes, the field has become much more dedicated to diversity and multicultural awareness. However, these changes have not yet instilled a strong ethos of cultural humility.

Changes in the Self-Report of Bias in Society and Psychology

The multicultural movement in psychology exemplifies the fact that events and trends in the broader American society impact trends within the field of psychology. Racial and cultural bias trends have been no different. Explicit forms of racism and bias have become less socially acceptable (Chong, 1991; Gitlin, 1987; Levy, 1992; Williams, 1987), which has caused a decrease in the self-report of racist ideas (Brigham, 1972; Karlins et al., 1969; Maykovich, 1971, 1972; Schuman et al., 1997). However, because of continuing racial group disparities and the lived experiences of BIPOC, it became clear to social psychology researchers that racism needed to be studied in new and creative ways that relied less on self-report (Crosby et al., 1980; Gaertner & Dovidio, 1977; Jones & Sigall, 1972; Nisbett & Wilson, 1977). Because of these efforts, ample evidence exists that bias against marginalized groups continues today, but largely
in more implicit and covert forms (Solórzano & Yosso, 2002; Sue, 2010). Furthermore, evidence exists that these attitudes can manifest themselves in individual judgments and behavior (Dovidio et al., 1997; Dovidio et al., 2002; Fazio & Hilden, 2001; Fazio et al., 1995; Sekaquaptewa et al., 2003, Experiments 1 and 2; McConnell & Leibold, 2001). Unfortunately, research suggests that these biases and their related behaviors exist within the field of psychology as well (Becker et al., 2003; Carson et al., 2014; Cohen et al., 2003; Cranford, 1999; Kugelman, 2018).

**Cultural Biases of Mental Health Professionals**

While most mental health professionals do not display explicit forms of racial or cultural bias, they are not immune to more implicit forms of bias. Implicit bias can show up in interpersonal behavior (Greenwald & Krieger, 2006). The biased actions of MHPs have been highlighted by an abundance of research across time (Becker et al., 2003; Carson et al., 2014; Cohen et al., 2003; Cranford, 1999; Guthrie, 1976; Kugelman, 2018). These studies have included a wide range of research designs, from vignette studies to studies conducted in naturalistic treatment settings. While some studies have failed to show bias, the majority have not. Bias has been shown in both clinical evaluations (such as diagnosis and treatment recommendations) and psychotherapists’ responses (such as reactions or perceptions) to their clients.

**Diagnosis**

Studies have outlined that psychotherapists’ racial bias impacts both overdiagnosis and underdiagnosis of BIPOC (Becker et al., 2003; Cranford, 1999; Kaden, 2010). One study conducted in a naturalistic setting found that even after controlling for severity of self-reported eating disorder symptoms, Latinx and Native American clients were still significantly less likely
than White clients to be referred for eating disorder treatment (Becker et al., 2003). Another study found that a hypothetical African American client with symptoms of schizophrenia was less likely than a hypothetical White/European American client to be diagnosed with schizophrenia by psychologists and psychology interns (Cranford, 1999). Underdiagnosis is problematic because it may lead to consequences such as less specialized treatment, or even a complete lack of treatment.

Overdiagnosis is also a problem within the field. For instance, one study found that when a hypothetical African American client presented with mood disorder symptoms, they were more likely than a hypothetical White client to be diagnosed with schizophrenia (Cranford, 1999). Another vignette study by Kaden (2010) found that psychotherapists were more likely to diagnose an African American teen with conduct disorder than they were to diagnose a White teen. In another vignette study by Rieffel (2006), psychotherapists were given either a White client or a Latinx client with the same characteristics. The researchers found that only 38% of the psychotherapists assessing the Latinx client accurately diagnosed them. In contrast, 71% of the psychotherapists diagnosing the White client were accurate. Overdiagnosis is problematic because it may lead to consequences such as stigmatization, over-exaggeration of symptoms, or even improper care.

**Treatment Recommendations**

In addition to differences in diagnosis, racial bias has been shown in treatment recommendations (Cohen et al., 2003; Gordon et al., 2006; Kugelmass, 2018). One study conducted by Kugelmass (2018) showed that White mental health care professionals were more likely to offer to speak on the phone or state that they had appointment availabilities to White clients. The same was not true of the Black professionals, who seemed to offer a phone call or
mention appointment availabilities at similar rates across clients. Another study conducted by Kugelmass (2016) found a similar trend when clients left voicemails seeking appointments. The study found that Black middle-class individuals were less likely than White middle-class individuals to be offered an appointment. Interestingly, Black and White working-class individuals were offered appointments at similar rates, but as a combined group, were three times less likely to be offered appointments than their middle-class counterparts. Another study found that dementia patients diagnosed with depression were more likely to be prescribed antidepressants if they were White (Cohen et al., 2003). As may be obvious, inaccurate or insufficient treatment recommendations could negatively impact clients of color in need of services and should be cause for concern.

**Psychotherapist Reactions**

Studies have also shown bias in counselor reaction to and perceptions of clients. For instance, Blanco (2010) used vignettes to examine whether clinical detection of intimate partner violence (IPV) differed based on the race of the client. The authors found that psychotherapists were significantly more likely to identify IPV for the White clients than for the Black clients. In another vignette study by Rieffel (2006), psychotherapists were given either a White client or a Latinx client with the same characteristics. The psychotherapists were then asked to determine if the client might be malingering. The researchers found that when psychotherapists stated that they suspected the Latinx client to be malingering, they portrayed more confidence in this judgment than those that made the same judgment with the White client. Studies such as these raise concern about the ways that stereotypes and biases about different racial/ethnic groups might impact psychotherapist’s reactions and perceptions of their clients. However, it is human nature to assume that such biases characterize the work of other people, not one’s own practice.
(Alicke et al., 2005; Gilovich et al., 2005; Pronin et al., 2002), with many psychologists needing cultural humility to recognize that consistent evidences of bias need to inform their own practice (Hook et al., 2013).

**Psychological Explanations for Mental Health Professionals’ Cultural Biases and Lack of Cultural Humility**

Psychological research and social psychological theories have sought to explore modern day cultural bias and its impact on individual behavior. Because of the many levels of investigation on this topic, Duckitt (1992) proposed a framework that integrates and connects the differing areas of research. His framework includes four causal processes of cultural biases: internal psychological processes, social and intergroup dynamics, social transmission, and individual differences. Given that social transmission processes are ubiquitous, the following sections will focus on the internal psychological processes and social and intergroup dynamics. Furthermore, because this review seeks to highlight the biases of psychotherapists as a group, and it will also exclude a discussion of individual differences.

**Internal Psychological Processes**

Although it is likely that most MHPs likely do not intend to disadvantage clients based on culture, all humans experience internal psychological processes, including some specific cognitive biases, that may impact their perceptions of other people (Alicke et al., 2005; Anthony et al., 1992; Gilovich et al., 2005; Jacoby-Senghor et al., 2015; Nisbett & Wilson, 1977; Oeberst et al., 2020; Paris et al., 1972; Pronin et al., 2002). Potential cognitive biases that may impact MHPs’ differential treatment of clients include implicit bias, ingroup bias, fundamental attribution error, ultimate attribution error, bias blindspot, continued influence effect, mere exposure effect, empathy gap, outgroup homogeneity bias, and just world phenomenon. While
this list of biases might not be comprehensive, it covers several likely cognitive processes contributing to MHPs’ cultural biases and lack of cultural humility.

**Implicit Bias.** The cognitive revolution in psychology, promoted by such scholars as Nisbett and Wilson (1977), confirmed that because humans are not always consciously aware of their mental processes, self-report measures may be inadequate and unreliable. Specifically, people tend to underreport their own biases (Alicke et al., 2005; Gilovich et al., 2005; Pronin et al., 2002). Over time, researchers sought to account for problems with self-reported data by developing measurements that did not rely exclusively on participants’ awareness. Particularly relevant to this dissertation, Greenwald and Banaji introduced the idea of *implicit bias* in 1995. Implicit bias refers to preferential responding that occurs outside of human awareness. To measure this phenomenon, Greenwald and colleagues (1998) created an Implicit Association Test (IAT), and since that time research on the topic exploded with over 10,000 articles citing the IAT since its creation (Kurdi et al., 2019).

Many research studies have confirmed an association between implicit bias and actual behavior (Agerström & Rooth, 2011; Amodio & Devine, 2006; Greenwald et al., 2009; Kurdi et al., 2019; Oswald et al., 2013). For instance, Greenwald et al. (2009) found that the IAT better predicts Black-White interracial behavior than self-reported explicit bias. This finding helps to explain why MHPs with no conscious awareness of bias might still act in biased and therefore, not culturally humble ways. Many of the biases outlined in this section often occur in the form of implicit bias. Furthermore, research has shown that even individuals who are skilled in assessing bias of others, may be ignorant of their own biases (Ehrlinger et al., 2005; Pronin et al., 2002). This phenomenon is called the bias blind spot.
**Bias Blind Spot.** The bias blind spot is the cognitive bias that causes individuals to be aware of others’ biases, but not their own (Alicke et al., 2005; Gilovich et al., 2005; Pronin et al., 2002). This bias is particularly relevant to MHPs, who are often academically trained to spot cultural bias, but may have less practice in the process of personal introspection about their own biases. Pronin and colleagues (2002) showed that people overwhelmingly assume that they are above average in freedom from bias. Furthermore, Ehrlinger and colleagues (2005) conducted four studies that supported the bias blind spot. Two of these studies showed that individuals believe that their life experiences connected to a topic serve them to have a more accurate opinion on the topic, while they view opposers’ relevant life experiences to be a source of bias. The other two studies showed that individuals are quicker to assume that they experience bias in the abstract than when assessing real life situations.

The bias blind spot helps to explain why MHPs may be intellectually aware of the concept of bias, but not necessarily aware of their own cultural biases. For example, a self-described feminist psychotherapist might have an excellent ability to recognize and point out sexism in her client’s lives but may hold cultural biases of her own and less likely to recognize this bias in herself. This phenomenon helps us to understand how educated and well-meaning psychotherapists may still hold cultural biases that impact their work with clients from different cultures.

**Ingroup Bias.** One such cultural bias is the ingroup bias. Ingroup bias was described by sociologist William Sumner (1906). Sumner theorized between-group competition as a competition for survival, and ingroup cooperation as a protection against harm. A field study conducted by Sherif and colleagues (1961) backed Sumner’s theory. This widely cited study placed boys in a summer camp into two groups, neither of which knew about the existence of the
other group. Upon finding out about the other group at the end of the eight days, the groups immediately began name-calling and commenting derogatorily about the other group (Sherif et al., 1961). The boys were then tasked with competing against one another, and ingroup bias was observed, with the boys favoring their own group members over outgroup individuals. Since this study, many more studies with varied designs have supported the existence of ingroup bias (Jacoby-Senghor et al., 2015; Oeberst et al., 2020; Paris et al., 1972).

This psychological process is at the very root of cultural humility and its opposite: The more psychotherapists see their clients in terms of their own cultural norms, the less likely they are to seek out and take seriously the perspectives and experiences of their clients. For example, a White American psychotherapist who values individualism and holds ingroup bias, may struggle to value and understand the perspective of a recent immigrant from Samoa who interacts with their family in a collectivistic way. In fact, the psychotherapist might be tempted to label this culturally relevant coping mechanism as “codependent.” Situations such as this highlight the importance of psychotherapists recognizing, understanding, and valuing cultural perspectives that differ from their own.

**Empathy Gap.** One of the key ways that psychotherapists come to understand their client’s perspectives is through empathy. Unfortunately, researchers have found that people have a difficult time empathizing with, helping, and even valuing the lives of outgroup members as much as they do with ingroup members (Gaertner et al., 1982; Kunstman & Plant, 2008; Pratto & Glasford, 2008; Saucer et al., 2005). This phenomenon is called the empathy gap. Research has shown this phenomenon can impact behaviors, and not just emotions. For instance, one study found that the more prejudiced individuals are, the more difficult it is for them to understand outgroup members’ emotional states and react appropriately (Gutsell & Inzlicht, 2012).
The empirical evidence for the empathy gap suggests that while mental health professionals are trained to respond in empathic ways, this skill may come easiest with members of their own cultural group. For example, a White American psychotherapist might have an easier time empathizing with and therefore detecting sadness and hopelessness in members of their own cultural group, and therefore struggle to diagnose a Native American client dealing with major depressive disorder. The empathy gap outlines a potential competency problem for MHPs and is likely one of the many ways that MHPs may struggle to relate to their clients from another group or culture.

**Outgroup Homogeneity Effect.** Another phenomenon that likely impacts MHPs ability to connect with and understand outgroup members is the outgroup homogeneity effect. The outgroup homogeneity bias is the tendency for individuals to perceive more variability among racial/ethnic ingroup members than members of other groups. This effect has been backed by much research (Jones et al., 1981; Park & Rothbart, 1982; Quattrone & Jones, 1980; Wilder, 1984). In fact, a meta-analysis of fifteen studies showed that White individuals have a harder time differentiating Black individuals’ faces than they do White individuals’ faces. Black individuals’ biases were also tested, but the effect was smaller (Anthony et al., 1992).

This effect may influence MHPs as they interact with clients of a different race or ethnicity than their own. This may specifically be a problem when MHPs are aware of multicultural research that focuses on generalities. While this research is important and relevant, unless used in a thoughtful and skillful manner, it may at times lead practitioners to hold stereotypes and generalizations that are unhelpful about their BIPOC clients that is unhelpful. The outgroup homogeneity effect helps to highlight the importance of cultural humility in the process of truly understanding the unique abilities and struggles of each client.
**Fundamental Attribution Error.** Another bias that may hinder the use of cultural humility in assessing outgroup members’ experiences, is the fundamental attribution error (also known as the correspondence bias). This bias causes individuals to believe that others’ behavior has more to do with fundamental personality traits, regardless of evidence about relevant situational factors. This phenomenon was first noted by Fritz Heider in 1958 and then explicated by Lee Ross in 1977. The phenomenon was backed up by a groundbreaking study by Edward Jones and Victor Harris (1967) in which they found that listeners assumed that individuals sharing an opinion were sharing their true opinions, even though the listeners were informed beforehand that the researchers instructed the individuals to share these opinions. Since these findings, studies have continued to confirm the widespread nature of fundamental attribution error (Jones et al., 1979; Snyder & Jones, 1974; Tukachinsky, 2020).

The fundamental attribution error may impact MHP thinking about racial and ethnic minority behavior that must be understood culturally and/or within the context of oppression. For instance, given the past and current racial inequality within the American justice system, many African Americans have a painful relationship with law enforcement, which may at times lead to an antagonistic view of the police. This may impact an African American teenager referred to treatment for consistent “bad behavior” in the form of “disrespect” toward police officers. Given the research on fundamental attribution error, even if an MHP is aware of these structural and historical factors, they may still attribute this “disrespectful behavior” to a personality flaw, or even a personality disorder such as Oppositional Defiant Disorder, rather than behavior based on an oppressive context. The fundamental attribution error highlights the ways that psychotherapists might see individuals from other contexts in negative or pathological ways more consistently than individuals within their own cultural context.
**Ultimate Attribution Error.** Similar to the fundamental attribution error is the ultimate attribution error (Pettigrew, 1979). The ultimate attribution error is the cognitive distortion of attributing bad behavior of an outgroup member to dispositional factors, and good behavior of outgroup members to a fluke, luck, high motivation or situational factors (Pettigrew, 1979).

It should be noted that while research supports the ultimate attribution error, an early review showed that research was not overwhelming (Hewstone, 1990). Still, after reviewing the literature, Hewstone (1990) wrote that the evidence seems to point to the ultimate attribution error contributing to stereotyping, intergroup hostility, and that it may help to form the basis of the idea that outgroup differences are due to biology.

Similar to the fundamental attribution error, the ultimate attribution error helps us to understand why some MHPs might hold biases against clients that are not part of their “ingroup,” even after experiencing the positive attributes of these clients. For instance, if the African American teenager described previously were to mention that they helped a woman recover her stolen purse from a robber, the psychotherapist might understand this by thinking that “anyone would do that” in the situation the teenager found themselves in. The ultimate attribution error helps to describe the process by which psychotherapists generate negative cultural biases.

**Just World Phenomenon.** As this paper has outlined, there are many misguided ways to understand the behaviors and experiences of outgroup members. One such outgroup includes individuals in less fortunate situations than oneself. Research has shown that it is common for individuals in privileged situations to believe that others in less privileged situations must have brought these circumstances on themselves in some way (Lerner, 1970; Lerner, 1965; Lerner & Miller, 1978; Lerner & Simmons, 1966). This phenomenon is entitled the just world
phenomenon (Lerner, 1980). This thought process helps the individual feel that the world is safe, and that bad things happening to other people won’t happen to them.

It is possible that this phenomenon occurs for MHPs as they notice which cultural groups have experienced greater hardships than their own and try to make sense of these trends in a way that allows them to feel safe. For example, a White American psychotherapist who was raised in a stable financial situation and has continued to experience financial stability during their adulthood may struggle to understand that the desperate financial situation of a recent refugee from the Democratic Republic of Congo. They might even believe (potentially implicitly) that these financial hardships are self-inflicted and could have been avoided with greater effort—an attitude that would likely get in the way of a positive therapeutic alliance. The just world phenomenon likely keeps MHPs from recognizing the reality of certain clients’ situations, and therefore may create an obstacle in the therapeutic relationship and clinical strategies. Unfortunately, research has shown that misinformation such as a belief in a just world may persist even in the face of discounting evidence.

**Continued Influence Effect.** The continued influence effect is the phenomenon of misinformation continuing to influence a person’s thinking, even after it has been disproven to the individual. For instance, the groundbreaking study by Ross and colleagues (1975) looked at participants’ perceptions of their performance on a task after having been given false feedback. The participants were informed of the falsehood of the feedback, and yet, these initial impressions of their performance remained. Decades of research have continued to back this phenomenon (Carretta & Moreland, 1983; Chan et al., 2017; Ecker et al., 2011; Johnson & Seifert, 1994; Lewandowsky et al., 2017; Lewandowsky et al., 2012; Rich & Zaragoza, 2016;
Ross et al., 1975; Schwarz et al., 2016; Wilkes & Leatherbarrow, 1988; Wyer & Budesheim, 1987).

This phenomenon can help to explain why cultural bias persists for MHPs, even after they are exposed to the strengths of other cultures and worldviews that disprove these biases. For instance, if an MHP holds stereotypes about Chinese individuals, and begins work with a Chinese individual who disproves these stereotypes both in their personal actions and their description of their culture, the continued influence effect can help to explain why these stereotypes persist. This phenomenon helps us to understand continued cultural bias in the minds of MHPs, and the need for cultural humility during mental health treatment. Luckily, while new information may not be enough to combat cultural bias, research has shown that mere exposure to other cultures may have an impact on individuals’ perspectives about those cultures.

**Mere Exposure Effect.** The mere exposure effect occurs when an individual is continually exposed to a stimulus, and without reinforcement, the repeated encounters improve their perspective of the stimulus (Zajonc, 1968). An early meta-analysis of the research supported the mere exposure effect as a valid phenomenon (Bornstein, 1989). Furthermore, research has shown that the mere exposure effect may influence White individuals’ perceptions of individuals from other races (Zebrowitz et al., 2008).

Given these findings, it is likely that the mere exposure effect impacts the level of cultural bias held by MHPs, especially for those who have little experience interacting with individuals from different races or ethnicities. However, the phenomenon also suggests that it is important for psychotherapists to seek out opportunities to interact with and get to know individuals from different cultures in order to gain respect for these individuals.
Social and Group Dynamics

There are many theories suggesting why humans repeat the internal psychological processes that retain cognitive biases that impact cultural bias. Among those relevant to this discussion include the Social Identity Theory, Realistic Group Conflict Theory, Optimal Distinctiveness Theory, Aversive Racism, and the Contact Hypothesis.

Social Identity Theory. The Social Identity Theory states that individuals experience ingroup members as part of their personal self-concept, and therefore prefer to see ingroup members more favorably (Tajfel et al., 1979). This is because negative views of one’s ingroup member may lead to negative views of oneself, while negative views of an outgroup member would be less likely to impact emotions about oneself. This theory helps to explain why MPHs might more easily see clients from within their culture more favorably, and clients from different cultures less favorably.

Realistic Group Conflict Theory. Not only do individuals experience ingroup members as part of their self-concept, they can also experience outgroup members as threats. This process is described in the Realistic Group Conflict Theory. This theory states that individuals tend to compete with members of other groups for symbolic (political power) or real (job security) resources (Campbell, 1965). This theory outlines why individuals in different groups struggle to relate to and interact peacefully with members of other groups and may also explain why MHPs may hold certain implicit cultural biases toward clients from outgroups. For instance, given the current often hostile political climate around immigration, certain MHPs may hold bias against a client who is a recent immigrant that has secured a position at a competitive company, as this may be experienced subconsciously as a threat to the psychotherapist’s job security.
Optimal Distinctiveness Theory. The Optimal Distinctiveness theory may be more applicable to MHPs than the last two theories, given that this theory helps to explain individuals that may hold cultural openness as a value, but not always act in line with it. This theory states that individuals have two competing social needs: the need for inclusion and the need for differentiation (Brewer, 1991). The theory states that individuals attempt to fill these needs by becoming part of groups that are inclusive of outgroup members to a degree, but also give them a sense of differentiating from those outside the group. Unfortunately, as outlined earlier in the paper, the field of psychology has a history of operating in this way—allowing for White Western psychologists to have the sense of being inclusive, but still different from BIPOC individuals and individuals from other cultures. This trend has impacted the level of cultural humility present within the field, psychological theories, and therefore in the therapy office.

Aversive Racism. Similar to the Optimal Distinctiveness Theory is the theory of Aversive Racism. This theory states that certain individuals who profess egalitarian views avoid contact or interaction with racial or ethnic minorities (Gaertner & Dovidio, 1986). The theory suggests that this is a subconscious phenomenon, and that the individuals may be unaware of their avoidant behaviors. This theory may help to further explain the lack of interaction between individuals of different races or ethnicities and may help to explain the continuation of cultural bias for MHPs. Aversive racism may even play out in subtle ways in which MHPs interact with and form relationships with different clients.

Contact Hypothesis. A final theory that helps to explain MHP cultural biases is the Contact Hypothesis. This hypothesis states that the more contact groups have with one another, the more they cooperate with each other (Allport, 1954). This is a hopeful solution to cultural bias but given the rates of segregated living that occur in the U.S., it is likely that individuals
most often interact with members of their own culture and, therefore, miss out on these opportunities for learning to interact peacefully and respectfully. This likely impacts MHPs whose social and professional networks are mostly inclusive of individuals from their own cultural group. Given the Contact Hypothesis, it may initially be more difficult for psychotherapists in this situation to work with clients from other cultural groups in unbiased ways. This is why multicultural competence is such an important aspect of MHP training.

**Psychotherapist Multicultural Competence: A Solution to Innate Cognitive and Social Cultural Biases**

*Multicultural Competencies in Psychotherapy*

Given the many biases described in the previous sections, it is vital that MHPs receive multicultural training in order to gain competence in working with a variety of clients. This may be particularly important for White counselors as there has been evidence to suggest that White MHPs may experience lower levels of multicultural competence than BIPOC MHPs (Holcomb-McCoy & Myers, 1999). These findings align with the history of psychology and its centralization of White Western culture (Guthrie, 1976). Furthermore, Richardson and Molinaro (1996) have suggested that White MHP self-awareness is a necessary ingredient for multicultural competence. It seems that understanding one’s implicit cultural biases is an important first step in multicultural competence and may be particularly important for White MHPs.

Thankfully, since the revolutionary work of Derald Wing Sue et al. (1982) and so many others, the field of psychology has begun to take multicultural competence seriously. The concept of Multicultural Competencies (MCCs) was first operationalized by Sue and colleagues in 1992, and then further clarified in a later article by Arredondo and colleagues in 1996. These
articles outlined three areas of MCCs which included attitudes/beliefs, knowledge, and skills. These standards have become a foundation for the entire field of multicultural psychology.

Research has supported the importance of multicultural competence in therapy. In fact, a recent meta-analysis showed that client-perceived cultural competence of the psychotherapist correlates strongly ($r = 0.38$) with treatment outcomes (Soto et al., 2018). Another recent meta-analysis found correlations between client-perceived multicultural competence of the psychotherapist and therapy process ($r = .75$) and outcome ($r = .29$; Tao et al., 2015). Furthermore, an extensive meta-analysis on the topic has shown that culturally adapted interventions are typically more effective than nonculturally specific interventions ($d = 0.35$ after accounting for publication bias; Soto et al., 2018). Similarly, research has shown that the more treatment is tailored to a client’s individual characteristics, the more likely the client is to be engaged and successful in therapy (Smith & Trimble, 2016).

Much research has supported the benefits of psychotherapists acquiring MCCs and has shown that psychotherapists can improve in their MCCs (Smith et al., 2006; Smith & Trimble, 2016). MCCs include multiple subcomponents within the three broad areas of awareness, knowledge, and skills. Given that breadth, it may be useful to conduct a narrower evaluation of individual components of the MCCs. In recent years, one of the components that has received substantial attention is cultural humility (Hook et al., 2013; Wright, 2019).

**Cultural Humility in Psychotherapy**

A key aspect of multicultural competence involves psychotherapists’ willingness to obtain additional knowledge, skills, and awareness relevant to clients’ cultural experiences and values. Practicing cultural humility during psychotherapy means that the psychotherapist proactively seeks to understand their clients’ cultural experiences and aligns therapy with these
experiences. Hook and colleagues (2013) describe this process as being “other-oriented (or open to the other) in relation to aspects of cultural identity” (p. 354). For instance, a White psychotherapist with an individualistic worldview may need to ask questions and be open to a Chinese American client’s collectivist worldview. If culturally humble, this psychotherapist might come to understand the strengths of this collectivist way of functioning instead of imposing their individualistic perspective on the client. As mentioned earlier, without this culturally humble approach, this psychotherapist might risk labeling the client “codependent” or see them as lacking personal autonomy or empowerment. However, if culturally humble, the psychotherapist will instead help to support the client in an empowered and connected way of living within their collectivist culture.

To further investigate the importance and impact of this construct, Hook and colleagues (2013) created a client-report measure of cultural humility. It should be noted that a measure based on psychotherapist-report was created by Gonzalez and colleagues in 2021, but because of the recency of its creation, research is very limited, and Hook et al.’s (2013) measure was instead used to operationalize cultural humility within the current study. Since the creation of Hook’s measure, research have explored the association between cultural humility and client experiences and outcomes. Most of the research investigating this association has been correlational and has used current and former clients’ evaluations of psychotherapist cultural humility along with other aspects of their therapy experiences. Research has shown that psychotherapist cultural humility positively correlates with the formation of a strong working alliance between psychotherapist and client (Hook et al., 2013; Wright, 2019). However, one study found that positive regard, empathy and congruence mediated this relationship (Wright, 2019). Furthermore, research has also shown a positive correlation between cultural humility and good therapy outcomes (Hook et
al., 2013; Owen et al., 2014; Owen et al., 2016). However, one study found this relationship to be mediated by a strong working alliance (Hook et al., 2013). The overall trend of the current research provides evidence that cultural humility is an important MCC, and that further research on the topic would likely benefit the broader understanding of MCCs.

Given the focus on multicultural competencies within the field, it is important for research on the varying aspects of MCCs to continue. Furthermore, we must work to better understand their unique impacts on therapy clients’ experiences and outcomes. Given that studies have already been conducted on the association between cultural humility and client experiences and outcomes, the current study seeks to synthesize these findings in the form of a meta-analysis. It is the authors’ hope that this meta-analysis will contribute to the broader understanding of the construct of cultural humility and its impact.
References


Guthrie, R. V. (1976). *Even the rat was white: A historical view of psychology* (Classical Ed.). Pearson.


https://doi.org/10.4324/9780203781159


https://doi.org/10.1002/ejsp.2420200404


https://doi.org/10.1002/j.1556-6676.1999.tb02452.x


APPENDIX B

Cultural Humility Scale

**Items**

**DIRECTIONS:** There are several different aspects of one’s cultural background that may be important to a person, including (but not limited to) race, ethnicity, nationality, gender, age, sexual orientation, religion, disability, socioeconomic status, and size. Some things may be more central or important to one’s identity as a person, whereas other things may be less central or important.

Please identify the aspect of your cultural background that is most central or important to you:

__________________________

How important is this aspect of your cultural background?

<table>
<thead>
<tr>
<th>Not at all important</th>
<th>Somewhat important</th>
<th>Very important</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>5</td>
</tr>
</tbody>
</table>

If there is a 2nd aspect of your cultural background that is important to you, please list:

__________________________

How important is this aspect of your cultural background?

<table>
<thead>
<tr>
<th>Not at all important</th>
<th>Somewhat important</th>
<th>Very important</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>5</td>
</tr>
</tbody>
</table>

If there is a 3rd aspect of your cultural background that is important to you, please list:

__________________________

How important is this aspect of your cultural background?

<table>
<thead>
<tr>
<th>Not at all important</th>
<th>Somewhat important</th>
<th>Very important</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>5</td>
</tr>
</tbody>
</table>

*PsycTESTS™ is a database of the American Psychological Association*
Please think about your counselor. Using the scale below, please indicate the extent to which you agree or disagree with the following statements about your counselor.

<table>
<thead>
<tr>
<th>Regarding the core aspect(s) of my cultural background, my counselor . . .</th>
<th>Strongly Disagree (1)</th>
<th>Mildly Disagree (2)</th>
<th>Neutral (3)</th>
<th>Mildly Agree (4)</th>
<th>Strongly Agree (5)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Is respectful.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>2. Is open to explore.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>3. Assumes he/she already knows a lot.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>4. Is considerate.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>5. Is genuinely interested in learning more.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>6. Acts superior.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>7. Is open to seeing things from my perspective.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>8. Makes assumptions about me.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>9. Is open-minded.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>10. Is a know-it-all.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>11. Thinks he/she understands more than he/she actually does.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>12. Asks questions when he/she is uncertain.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
</tbody>
</table>
APPENDIX C

Coding Sheet

A. **Short Description**

a. For Authors with only one article included in our study: Last name of first author and the last two digits of year (e.g., Belavich 98)
b. For authors that have more than one article included in our study that have also been published in the same year: Last Name + last two digits of year + Journal abbreviation in CAPS (ex. Koenig 99 AJG)
c. For authors with multiple studies in one article: Last name + last two digits of year + lower case letter (in sequential order) (ex. Levin 99a, Levin 99b....etc.)
d. For articles with multiple/different samples in a single study: Last name + last two digits of year + lowercase letter + number (in sequential order) (ex. Levin 98a1, Levin 98a2....etc.)

B. **APA style citation**  MAKE SURE ALL information appears (full page numbers, etc.)
APA FORMAT. You can obtain this information by clicking on the “ symbol in scholar.google.com

C. **Year of publication**

D. blank

E. blank

F. **Published**  0 = not published (dissertation, presentation)
1= published

G. blank

H. blank

I. **Field** (audience for article, can look at journal, authors’ departments, etc.)

1. psychology (includes counseling psychology)
2. sociology
3. medicine (includes psychiatry)
4. nursing
5. social work
6. education
7. religion
8. family and human development
9. multicultural or ethnic studies (even if also one of the above, i.e., JMCD)
10. counseling

J. **Coder(s):** Enter your names

K. **Geographic location** (of data collection = where people completed the measures)

Leave blank if no information or if not in North America

0 = multi-site (many data collection locations, such as a national survey)
1 = South (Virginia to Louisiana to Kentucky)
2 = South west (Texas to Arizona)
3 = West coast (California to Washington)
4 = Western states (Nevada to Colorado)
5 = Central states (Oklahoma to Dakotas, Kansas to Ohio)
6 = Eastern states (Maine to Washington DC)
10 = Canada

L. **Type of Sample of Psychotherapists/Professionals** (of the participants included in this row/effect size) **over 70%**

1. psychologists (including counseling psychologists)
2. psychiatrists
3. physicians (medical doctors treating mental health patients)
4. nurses
5. social workers
6. educators, school counselors, school psychologists (do NOT code this article)
7. religious clergy/leaders (e.g., pastoral counseling) (do NOT code this article)
8. marriage & family psychotherapists
9. other professionals (e.g., diversity trainers, academics) (do NOT code this article)
10. counselors (LMHC, masters level)
11. psychology students or interns
12. psychiatry students (residents or interns)
13. medical students
14. student nurses
15. social worker students
16. student educators, school counselors, school psychologists (do NOT code this article)
17. student religious personnel (do NOT code this article)
18. marriage & family psychotherapist students
19. other students (do NOT code this article)
20. counseling students
21. psychologists & psychology students
22. psychiatrists & medical students
23. physicians & medical students
24. nurses & nursing students
25. social workers & MSW students
26. educators, school counselors, school psychologists & students (do NOT code this article)
27. religious clergy/leaders & students (do NOT code this article)
28. marriage & family psychotherapists & students
29. other professionals & students (do NOT code this article)
30. counselors & students
31. MULTIPLE of the above (no clear majority > 70%)

M. **Mean Age of Psychotherapists/Professionals** (of the participants included in this row/effect size)

N. **Is the Mean age an educated guess or based on a given range?**

1. actual value reported
2. median of range given  (use only when the range is narrow - or you can weight multiple categories by n using spreadsheet)
3. guess based on sample description (guidelines but not the only options)
   20 = undergraduates
   24 = masters students
   27 = doctoral students
   29 = doctoral interns

O. **Percent female of Psychotherapists/Professionals** (of the participants included in this row/effect size)

P. **Educational Level = mean years of education of Psychotherapists/Professionals**
   If not reported, the following can be used as estimates:

   17= 1st year masters students
   18= 2nd year masters students or masters graduates (MFT, MSW, LPC, etc.)
   19= 1st year doctoral students
   20= 2nd year doctoral students
   21= 3rd year doctoral students or doctoral interns
   22= doctoral graduates, psychologists, physicians, etc.

Q. **Reported Multicultural Training of psychotherapists or levels of multicultural training**

1= not reported
2= Limit training, one class or workshop or described as somewhat experienced with clients evaluated
3= Received specific training/experience for the population study; two to three classes/workshop
4= Received high quality detailed or extensive training; four or more classes/workshops

R. **Ethnicity Reported of Psychotherapists/Professionals**

1. no
2. yes

S. **Percent White American** (of the participants in this row/effect size– same for all below)
T. **Percent African American**
U. **Percent Hispanic/Latinx American**
V. **Percent Asian American**
W. **Percent Native American**
X. **Percent “other” American** (race not specified or not included in one of the above)
Y. **Percent White Canadians** (Canadians of European origin)
Z. **Percent Canadians of Color** (First Nations and Canadians of Asian, African, Latinx, Pacific Island origin)

AA. **Type of Sample of Clients/Patients** (of the participants included in this row/effect size)
(Do not code this for client vignette studies)

*Note: Although the clients may have a variety of conditions, they should be treated for a mental health condition found in the DSM (in addition to anything else they may also have). Do not code if no mental health/emotional condition.*

1. hospital, physical conditions
2. outpatient treatment, physical conditions (physically sick)
3. nursing home
4. hospital, mental illness (include assisted living for SPMI)
5. outpatient mental health clinic/university counseling center
6. grieving/bereaved (non-clinical) e.g., widowed
7. church or religious group
8. normal adult community members (non-clinical, non-church)
9. university students (non-clients)
10. high school/jr. high
11. children
12. HIV positive
13. caregivers of patients/elderly (family of patients)
14. recently divorced
15. new mothers/teenage mothers
16. substance abuse
17. prisoners
18. unemployed
19. abuse survivors
20. LGBTQ+
99. UNSPECIFIED or mixed/multiple groups above

AB.  **Mean Age of Clients/Patients** (of the participants included in this row/effect size)
(Do not code this for client vignette studies)

AC.  **Is the Mean age an educated guess of based on a given range?** (Do not code this for client vignette studies)

0 = actual value reported  
1 = median of range given  
2 = guess based on sample description  
(10 = if the children are about 5th grade; 20 = age guess for undergraduates)

*Note:* Be cautious about assigning a “guess” value for clients unless the verbal description is fairly clear about the age of clients. So it would be OK to guess an average age of 66 for “recent retirees,” but please do not guess an age for “elderly” since they could be 80 to over 100.

AD.  **Percent female of Clients/Patients** (of the participants included in this row/effect size)
(Do not code this for client vignette studies)

AE.  **Educational Level = mean years of education of Clients/Patients** (# of years of education)
(Do not code this for client vignette studies)

For example (can code outside of these ranges):

11= community sample or low SES
12= high school graduates
13= college freshmen
14= college sophomores
15= college juniors
16= college seniors and college graduates
17= 1st year masters students
18= 2nd year masters students or masters graduates (MFT, MSW, LPC, etc.)
19= 1st year doctoral students
20= 2nd year doctoral students
21= 3rd year doctoral students or doctoral interns
22= doctoral graduates, psychologists, physicians, etc.

AF.  **Socioeconomic Status of Clients/Patients** (Do not code this for client vignette studies)

*MUST include evidence beyond mention of education level*

1= lower class (mean income below poverty line of 25k, low education attainment)
2= lower middle-class (income about 45k, high school education equivalent)
3= middle class (income about 70k, college education equivalent, graduate
students)
4= upper middle class (mean income about 100k, professional career)
5= upper class

Leave blank if insufficient information

AG. Is the SES an educated guess or based on a given range? (Do not code this for client
vignette studies)
1= actual value reported
2= median of range or other information given
3= guess based on sample description

AH. Ethnicity Reported of Clients/Patients (Do not code this for client vignette studies)
0= no
1= yes

AI. Percent White American (of the participants included in this row/effect size – same for
all below)
AJ. Percent African American
AK. Percent Latinx American
AL. Percent Asian American
AM. Percent Native American
AN. Percent “other” American (race not specified or not included in one of the above)
AO. Percent White Canadians (Canadians of European origin)
AP. Percent Canadians of Color (First Nations and Canadians of Asian, African, Latinx,
Pacific Island origin)

AQ. Language use of Clients/Patients (Do not code this for client vignette studies)
1= clearly English proficient
2= some doubts about English proficiency (English is indicated as a second
language, etc.)
3= likely not proficient in English (recent immigrants from a non-English
speaking nation, translation/interpretation used, etc.)

AR. Treatment type
99 = no real clients (vignette study, therefore no actual treatment provided)
0= real clients but no intervention/treatment provided
1= individual psychology/counseling
2= individual substance abuse (do NOT code)
3= individual other (do NOT code)
<table>
<thead>
<tr>
<th>Code</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>group psychology/counseling</td>
</tr>
<tr>
<td>2</td>
<td>group substance abuse (do NOT code)</td>
</tr>
<tr>
<td>3</td>
<td>group other (do NOT code)</td>
</tr>
<tr>
<td>4</td>
<td>family psychology/counseling</td>
</tr>
<tr>
<td>5</td>
<td>family other (do NOT code)</td>
</tr>
<tr>
<td>6</td>
<td>community/contextual intervention targeting mental health</td>
</tr>
<tr>
<td>7</td>
<td>more than one of 1-9 above</td>
</tr>
<tr>
<td>8</td>
<td>culturally specific individual psychology/counseling</td>
</tr>
<tr>
<td>9</td>
<td>culturally specific individual substance abuse (do NOT code)</td>
</tr>
<tr>
<td>10</td>
<td>culturally specific individual other (do NOT code)</td>
</tr>
<tr>
<td>11</td>
<td>culturally specific group psychology/counseling</td>
</tr>
<tr>
<td>12</td>
<td>culturally specific group substance abuse (do NOT code)</td>
</tr>
<tr>
<td>13</td>
<td>culturally specific group other (do NOT code)</td>
</tr>
<tr>
<td>14</td>
<td>culturally specific family psychology/counseling</td>
</tr>
<tr>
<td>15</td>
<td>culturally specific family other (do NOT code)</td>
</tr>
<tr>
<td>16</td>
<td>culturally specific community/contextual intervention targeting mental health</td>
</tr>
<tr>
<td>17</td>
<td>more than one of 11-19 above</td>
</tr>
</tbody>
</table>

**AS. Design type (for how this EFFECT SIZE was measured, which may differ from overall study design)**

<table>
<thead>
<tr>
<th>Code</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>Retrospective survey, ex post facto (looking back to what happened in the past)</td>
</tr>
<tr>
<td>1</td>
<td>Cross-Sectional, evaluating the present time (one-time surveys involving current psychotherapists and/or clients)</td>
</tr>
<tr>
<td>2</td>
<td>Predictive, longitudinal (repeated data across time – no intervention conducted)</td>
</tr>
<tr>
<td>3</td>
<td>Archival analysis (using existing databases or clinic records)</td>
</tr>
<tr>
<td>4</td>
<td>Comparison groups (≥2 groups, such as different races, but no control group [group not receiving treatment])</td>
</tr>
<tr>
<td>5</td>
<td>Experimental (≥2 groups, with a control group that did not receive what the experimental group received)</td>
</tr>
<tr>
<td>6</td>
<td>Pre to post-test comparison (single group, over time, with intervention - but no control group)</td>
</tr>
<tr>
<td>7</td>
<td>Factor analytical (factor analysis) (rare)</td>
</tr>
<tr>
<td>8</td>
<td>Single subject design (data from single participants over time) (do NOT code)</td>
</tr>
<tr>
<td>9</td>
<td>Panel study (cohorts followed across time) (rare)</td>
</tr>
<tr>
<td>10</td>
<td>Analogue cross-sectional (no actual clients = either one rater gets multiple vignettes [comparing differences across vignettes rated by the same psychotherapist]; OR everyone gets the same exact vignette, with a correlation between the psychotherapist’s evaluation of a vignette/fake client and the psychotherapist’s personal level of bias/comfort/cultural humility)</td>
</tr>
</tbody>
</table>
| 11   | Analogue comparison groups (no actual clients = scenarios/vignettes about ≥2 races (e.g., psychotherapists look at different vignettes of different races [which
is not a control condition but rather a comparison between races]) (VIGNETTES COMPARING CLIENTS OF DIFFERENT RACES ARE 11s)
12= Analogue experimental (no actual clients = scenarios/vignettes given to ≥ 2 different groups determined by the researchers [on a factor other than client race], such as (example 1) all psychotherapists receive the same form but then one group receives different instructions than another group or (example 2) psychotherapists receive different forms and then in addition, one group reports their racial bias/comfort but the other does not.
13= Analogue pre-post comparison (no actual clients = scenarios/vignettes given >2 times to a single group of psychotherapists, with something that occurred between the assessment using vignettes/scenarios - but no control group)(rare)

(Remember, analogue studies present artificial settings or conditions [such as watching a video or reading a vignette and rating it]= not the psychotherapist’s own actual clients)

AT. Random Selection (external validity; participant recruitment; pay attention to use of the word random)

1=convenience sample (volunteers, those who are accessible)
2=random sample (or entire population) with severe attrition (>60% lost)
3=random sample from limited area (1-3 sites) or all from one site
4=random sample from large area (>4 sites or > 1 million population)
5=mixed random sample and convenience sample (some participants were randomly selected, but others were not)

AU. Random Assignment (most vignette studies will be 2 or 3)

1= no group divisions (often the case for correlational studies, with data in terms of r )
2= convenience assignment, self-selection to groups (e.g., if people showed up for treatment or not, that was a matter of client convenience/self-selection; or if the independent variable is client race, clients already are whatever race they are, they are not randomly assigned)
3= random assignment to groups (must use the word “random” or a synonym like “computer generated” to describe how participants received the different conditions).

Ask 2 questions: 1) Were there any group comparisons? 2) How were groups assigned?

Subsequent variables deal with RESULTS
one row per effect size extracted – then compute an aggregate as needed
AV. **Approach to topic = dependent variable type for THIS EFFECT SIZE**

111 = Psychotherapist judgments of client symptom severity, diagnosis, prognosis, or success in treatment *(clinical variables)*
112 = Psychotherapist judgments of client *non-clinical* behaviors/characteristics/impressions/experiences (e.g., likeability, attractiveness)
113 = blank
114 = Psychotherapist judgments of the type of treatment needed/recommended or whether there was an offer to treat the client (call back)
115 = Psychotherapist anxiety/comfort/bias level/microaggressions
116 = Therapy outcome measure (client change as a result of mental health treatment/intervention), including client perceptions of improvement or change due to treatment
117 = Client perceptions of therapeutic working alliance or relationship with psychotherapist
118 = Client satisfaction with treatment
119 = Multiple categories (more than one of the above)

AW. **Factor supposed to influence outcome = independent variable**

1 = Race of the client (White vs. People of Color)
2 = Race of the client (People of Color compared with one another)
3 = Race of the client (Mixed/multiple comparisons: both BIPOC & BIPOC and also BIPOC & White)
4 = Race of the psychotherapist
5 = Racial attitudes/beliefs of the psychotherapist (racial empathy/awareness/comfort/anxiety/ethnocentrism)
6 = Cultural humility of the psychotherapist (CHS Total; must be “cultural humility,” not synonyms for this one)
6.1 = CHS positive subscale
6.2 = CHS negative subscale

AX. **Race of client comparison (for this effect size row)**

0 = no racial comparison across clients (e.g., correlational study of psychotherapist bias; only psychotherapist race compared)
1 = African American/Black clients compared with White clients
2 = African American/Black clients compared with Asian American clients
3 = African American/Black clients compared with Latinx clients
4 = African American/Black clients compared with other clients of color (groups not listed above)
5 = Asian American clients compared with White clients
6 = Asian American clients compared with Latinx clients
7 = Asian American clients compared with other clients of color
8 = Latinx/Hispanic clients compared with White clients
9 = Latinx/Hispanic clients compared with other clients of color
AY. What type of effect size?

Remember—You can only have ONE 1, 4, or 6 (last row)

1 = This is the ONE effect size in this article (There is only one effect size in this article, and this is it.)

2 = This is one of >1 effect sizes inclusive of the whole sample
Multiple dependent measures administered, each measure is a 2 on a separate line i.e., one of several effect sizes computed with the entire sample

3 = This is one of >1 effect sizes for subgroup
This effect size pertains only to a subset of people (e.g., for women, reported separately from men), or results separated by race, age, psychotherapist type, or any other group attribute, not the entire sample.

4 = Aggregate of multiple effect sizes with all participants (computed average of only 2s) i.e., you need to average multiple effect sizes where more than one are reported for the entire sample.

6 = Aggregate of effect sizes for subgroups (computed average of 3s, weighted by n) i.e., you need to average multiple effect sizes where more than one are reported for different subgroups. Basically any time you need to weight by n in order to calculate the aggregate effect size. (women only, African Americans only) – use the “weighted by n” tab

*Anytime 3s and a 6 for the aggregate, the 6 should have the largest number of participants. If 2s differ slightly in N (usually be the same or very similar), report the largest number for the 4 agg.

AZ. Statistics (metric of the value reported in the manuscript)

1 = Zero order correlations (Pearson r, spearman rho, phi coefficient) (only two variables)
2 = Partial correlations, beta weights (regression), path coefficients (controlling/adjusted for other variables)
3 = ANOVA (F-tests), MANOVA (requires different calculation than ANOVA → use Wilson calc link)
4 = t-test
5 = Odds ratios
6 = Chi square
7 = Means & Standard Deviations, or Cohen’s $d$ (mean difference)
8 = ANCOVA/MANCOVA (analysis of covariance) (controlling/adjusted for other variables) - be sure to code covariates later
9 = P value only (or when ES=0 due to non-significance)
10 = Combination (particularly for aggregates)
11 = Percentages or frequency counts

BA. **Racial comparisons across psychotherapist/client**

0 = counselor and clients are of the SAME race

1 = white counselor vs. black clients (assuming comparison group is white clients)
2 = white counselor vs. latino clients (assuming comparison group is white clients)
3 = white counselor vs. asian clients (assuming comparison group is white clients)
4 = black counselor vs. white clients (assuming comparison group is black clients)
5 = black counselor vs. other clients (assuming comparison group is black clients)
6 = latinx counselor vs. white clients (assuming comparison group is latinx clients)
7 = latinx counselor vs. other clients (assuming comparison group is latinx clients)
8 = asian counselor vs. white clients (assuming comparison group is asian clients)
9 = asian counselor vs. other clients (assuming comparison group is asian clients)

10 = multiple or matrix comparisons (2x2 table; chi square value; or more than 2 racial groups at the same time; or unknown/many combinations of psychotherapist and client race)

11 = other group combinations NOT listed above or below

12 = white clients vs. black counselor (assuming comparison group is white counselor)
13 = white clients vs. latinx counselor (assuming comparison group is white counselor)
14 = white clients vs. asian counselor (assuming comparison group is white counselor)
15 = black clients vs. white counselor (assuming comparison group is black counselor)
16 = black clients vs. other counselor (assuming comparison group is black counselor)
17 = latinx clients vs. white counselor (assuming comparison group is latinx counselor)
18 = latinx clients vs. other counselor (assuming comparison group is latinx counselor)
19 = asian clients vs. white counselor (assuming comparison group is asian counselor)
20 = asian clients vs. other counselor (assuming comparison group is asian counselor)
BB. **Effect size $d$ when the independent variable is group/race/category differences**

For $d$, enter the value.
For all other types of group/category data (such as comparing outcomes/ratings across racial groups), convert the value to $d$ using the effect size software (i.e., for odds ratios, F values, etc.)

Positive values = beneficial effect or positive bias toward People of Color

Negative values = harmful effect or negative bias against People of Color

For aggregate effect sizes based on adding up subgroups (6), *weight the overall ES calculated by the N of each subgroup.*

BC. **Effect size when the independent variable is a continuous measure (e.g., scales/measures of psychotherapist bias)**

Code as correlation coefficient
For $r$ or standardized beta weights in a regression model, enter the value.
For all other types of continuous data, convert the value to $r$ using the effect size software.

Positive values = the stronger the psychotherapist positive beliefs toward People of Color, the better clients’ outcomes
or the more negative psychotherapists beliefs towards People of Color, the worse clients’ outcomes
(this is the expected association between psychotherapist actions and client experiences)

Negative values = the stronger the psychotherapist positive beliefs toward People of Color, the worse clients’ outcomes
or the more negative psychotherapists beliefs towards People of Color, the better clients’ outcomes
(unexpected association between psychotherapist actions and client experiences)

*Note:* do NOT code all correlation/regression coefficients in this column! Only insert data in this column when

BOTH the independent variable and the dependent variable are continuous data, with the independent variable being some indication of psychotherapist racial attitudes/bias

BD. **Statistical Controls**
LEAVE THIS BLANK if there are no statistical controls (i.e., if it is not a regression model, ANCOVA, partial correlation, or path)
Report the name of the variable(s) that are controlled for in the result reported.

**BE.** **Total sample size that this effect size is based on**

*Note:* Often different from the initial N reported due to attrition. This should be the N used in this specific analysis (for this effect size), such as found in the footnote of a table or in the text reporting degrees of freedom (N = df plus the # of groups)

For aggregates, use largest N if averaging 2s; use the summed N if combining 3s

**BF.** **Sample size, experimental group** (if any) (if grouped by race, use the N of people of color)

*Note:* When dividing total N arbitrarily and total N is odd, give the extra person to the experimental group.*

If a correlation, leave BF and BG blank and only give the total N in BE

**BG.** **Sample size, control/comparison group** (if any) (if grouped by race, use the N of whites or the second group of people of color listed in the data)

**BH.** **Name of measure of the outcome variable** (specific name and subscale)

*(what was supposed to have been impacted = DEPENDENT variable)*

If the scale is homemade, type in a brief description of what was measured (e.g., judgments of clients’ likelihood to succeed in therapy) and specify “homemade” = 99

**BI.** **Reliability coefficient of the dependent measure** (alpha/Cronbach’s alpha, or internal consistency coeff.)

**BJ.** **How was reliability determined?**

0= Actual value *with this sample* was reported

1= The reliability was based on ANOTHER data set (i.e., from the original study of that measure)

2= The reliability was based on averaging two or more reliabilities (aggregate or median)

3= (do not use this number to code)

4= Derived using the Spearman-Brown Prophecy Formula for shorter/lengthened measure

5= Reliability derived from extrapolating reliabilities from similar measures (rare)

**BK.** **Name of independent variable or other measure** (type in specific name and subscale)

*(What was supposed to make a difference, such as the psychotherapists’ attitudes/comfort with certain racial groups)*
Or specify the racial groups that were being compared
If the scale is homemade, type in a brief description of what was measured (e.g., attitudes toward clients)

For the Cultural Humility Scale codes, enter the Cronbach alpha/internal consistency coefficient here ONLY

BL. Leave blank

BM. Verbal descriptions and clarifications of this study (IF NECESSARY)
(use this column to denote interesting elements of the study)

BN. Clarifications of problematic coding issues for this study (use this column to tell us how you coded things that were unclear – what page you found your evidence to support unclear coding, etc.)

BO. Psychotherapist awareness variable (did psychotherapists know they were being evaluated across clients’ race?)

0 = No information can be deduced about whether the psychotherapists might have known that they were being observed as a function of client race

1 = evaluation by race was NOT known by the psychotherapist (psychotherapists were completely unaware that they were being observed across racial conditions) (archival studies or studies in which psychotherapists were unaware that they were being evaluated based on differential treatment of racial groups, not knowing that they would be compared across client race)

2 = psychotherapist aware of client race of only one group (each participant in the study rated a person or people from a single racial group. For example, a mailed survey with the race of one client was provided; the study evaluated psychotherapists who only responded to one race, not racial differences by the same respondent; differences across psychotherapists who evaluated different races from one another)

3 = psychotherapist aware of client race, evaluating/responding to 2 or more racial groups (each participant in the study rated/treated people from at least two different racial groups, and they were aware of clients’ race; for example, a mailed survey with ≥2 clients of different races described or rated by the same psychotherapist)

4 = psychotherapist is aware of the topic of race by completing a MEASURE of racial attitudes or racism (psychotherapists were directly asked about their racial attitudes, cultural humility, etc. Direct questions about attitudes about race or racial issues; usually this would heighten the awareness that this was a
multicultural study or a survey of their racial attitudes.) If both 4 and any other code option, code as 4.

5 = psychotherapists were aware that race was part of the study, but the article did not provide enough information to clearly code one of the above values (similar to numbers 2 and 3 above, except completely unclear about how much the psychotherapists were exposed to information about client race).

BP. Who completed the Cultural Humility Scale

1 = current client
2 = former client (retrospective recall)
3 = psychotherapist

BQ. Who completed the dependent/outcome measure for this Cultural Humility effect size

1 = current client
2 = former client (retrospective recall)
3 = psychotherapist
4 = observer/rater (neither psychotherapist nor client)

BR. Mean score on the Cultural Humility Scale - divided by the number of items

When authors report an overall average/mean for the CHS, they can use different formats for reporting:
- When they provide the overall/total average score and also the number of items (should usually be 12 unless a subscale), divide the average score by the number of items actually given to participants in this study.
- If they do not report the number of CHS items that they administered, assume 12 and divide by 12, but then make certain that the resulting value is reasonable (should be between 40 and 55)
- If they report an average CHS score between 2 and 5, they already divided by the number of items. Simply code that number.

BS. Standard deviation of the Cultural Humility Scale - divided by the number of items

Follow the same logic as with the mean score, keeping SD “as is” if item-level mean is already reported by authors, otherwise dividing by the number of items.