The Effectiveness of Digital Interventions for Enhancing Empathy in Adults:

A Meta-Analysis of Randomized Controlled Trials

Camille Y. Williams[1,2], Qiang Xie[1,2], Elvan Muratoglu[3], Simon B. Goldberg[1,2]

[1]Department of Counseling Psychology, University of Wisconsin – Madison, WI, USA

[2]Center for Healthy Minds, University of Wisconsin – Madison, WI, USA

[3]Michigan Medicine, University of Michigan – Ann Arbor, MI, USA

Author Note

Camille Y. Williams: https://orcid.org/0000-0002-9324-3014

Qiang Xie: https://orcid.org/0000-0003-3968-9629

Simon B. Goldberg: https://orcid.org/0000-0002-6888-0126

Correspondence concerning this article should be addressed to Camille Y. Williams, Department of Counseling Psychology, University of Wisconsin – Madison, 335 Education Building, 1000 Bascom Mall, Madison, WI, 53706, cywilliams@wisc.edu or Simon B. Goldberg, Department of Counseling Psychology, University of Wisconsin – Madison, 335 Education Building, 1000 Bascom Mall, Madison, WI, 53706, sbgoldberg@wisc.edu

Statements and Declarations

Funding

Conflicts of Interest

The authors have no relevant financial or non-financial interests to disclose.

Ethics approval

Ethics approval is not applicable.

Informed Consent

Informed consent is not applicable.

Preregistration

This study was preregistered through the Open Science Framework

(https://osf.io/b9ukd?view_only=b70bba762eee4c11ac3528ddd3e79b38).

Abstract

The current study examined the effectiveness of digital interventions for enhancing empathy in adults. We conducted a meta-analysis of randomized controlled trials (RCTs) of interventions compared with non-specific (i.e., not intended to be therapeutic) and active control conditions (i.e., intended to be therapeutic). Across 24 studies, including 26 comparisons ($n = 3,137$), the overall estimated effect size was statistically significant ($g = 0.19$, 95% CI [0.05, 0.32], $p = .006$). However, when accounting for publication bias using trim-and-fill analysis, the estimated effect size became smaller and non-significant ($g = 0.07$, 95% CI [-0.07, 0.22], $p = .322$). Studies with follow-up data were analyzed separately ($k = 6$), yielding a small effect ($g = 0.29$, 95% CI [0.07, 0.5], $p = .008$) which was robust to trim-and-fill analysis. No significant moderators were identified. An exploratory analysis of studies separated by control group type revealed a small positive effect for studies with non-specific control groups ($g = 0.22$, 95% CI [0.12, 0.33], $p < .001$) which was robust to trim-and-fill analysis. The results suggest there is potential for digital interventions to have a positive impact on enhancement of empathy in adults, although digital interventions do not appear to outperform active control conditions. For digital interventions to most fully deliver on their potential, they may require more advanced technology that can offer attuned interaction that mimics a human relationship, as well as offer feedback to participants.

Keywords: Empathy; smartphone; virtual reality; RCT; meta-analysis; systematic review

The Effectiveness of Digital Interventions for Enhancing Empathy in Adults: A Meta-Analysis of Randomized Controlled Trials

Broadly conceptualized as the ability to imagine another's perspective and subsequently have a vicarious emotional experience, empathy plays a critical role in human interaction (Cuff et al., 2016). It underpins healthy and fulfilling relationships with others and with oneself (De Paul & Guibert, 2008; Gibb & Abela, 2008). Researchers describe empathy as both a trait with a developmental trajectory as well as a state influenced by specific social contexts (Cuff et al., 2016). In addition, empathy can be categorized into cognitive and affective components. Cognitive empathy is the recognition of emotion in another (Gerdes et al., 2010) and can be understood as a top-down process that involves producing a working model of others' emotional states based on visual, auditory, and context clues and thus requires holding and employing information within the working memory (Decety & Meyer, 2008). In contrast to cognitive empathy, affective empathy is a bottom-up process of experiencing another's emotions oneself and responding with an appropriate emotion (Cuff et al., 2016). Empathy can also have a behavioral component (i.e., outward expressions of empathy; Overgaauw et al., 2014).

Empathy is important for individuals, relationships, and society. Evidence overwhelmingly suggests that increases in empathy lead to higher quality of life, healthy and positive relationships, and improved well-being (Morelli et al., 2015; Vinayak & Judge, 2018). On a societal level, empathy plays a powerful role in the welfare and progress of all communities and cultures (Decety & Jackson, 2004; Enke, 2023). Indeed, the enhancement of empathy has been a focus of many researchers in developmental and educational psychology because it has the power to shape responsible citizenship and societal change within the next generation (Krznaric, 2014; Rigg & van der Wal-Maris, 2020). Additionally, empathy in helping

professionals such as psychotherapists or physicians is associated with better outcomes for patients (Hojat et al., 2011; Watson et al., 2014). Numerous interventions have been created for fostering empathy in children, such as socio-emotional learning curriculums for preschool and elementary school classrooms (Rosenthal & Gatt, 2010; Schonert-Reichl et al., 2012). While considerable research has examined interventions that target empathy in the early stages of life, relatively less attention has been given to interventions that might enhance empathy in adults.

Interventions that have been studied for enhancing empathy in adults have taken various forms, the most common of which have consisted of in-person trainings that include games, role play, didactics (lectures), demonstrations and practice, interpersonal-process-recall videotape-feedback models, guided meditation, and micro-training (counseling micro-skills trainings; Lam et al., 2011; Teding van Berkhout & Malouff, 2016). Research evaluating these empathy trainings indicate that they are generally effective. Indeed, a meta-analysis conducted by Teding van Berkhout and Malouff (2016) examined the efficacy of empathy training, primarily focusing on traditional, non-digital interventions. The results showed a medium magnitude effect ($g =$ 0.63) favoring empathy training over controls. While it is established that traditional empathy trainings are effective, the efficacy, scalability, and accessibility of these interventions are limited by lack of immersive features and practical barriers such as cost, the need for physical transportation, and limited resource availability.

Digital interventions have the potential to enhance the efficacy, scalability and accessibility of empathy trainings. For the purpose of this meta-analysis, we define digital interventions broadly to encompass a variety of digital technologies including computers, mobile devices, online tools, virtual reality (VR), digital role-playing simulations, and video games, among others that are used with the aim of affecting an outcome or change. Digital interventions

may elicit higher engagement and result in more significant and longer-term impact than traditional trainings. For example, in a study that aimed to elicit empathy for the homeless, Herrera et al. (2018) found that participants in the VR perspective-taking task had more positive, longer-lasting attitudes toward the homeless and signed a petition supporting the homeless at a significantly higher rate than participants who completed a traditional perspective-taking task. It is likely that the immersive nature of VR, simulation, video game, and computer game platforms provide a more realistic, embodied experience of empathy (Boltz et al., 2015), potentially resulting in more potent impact. Additionally, these types of digital interventions may be more effective because they offer a space that is safe to make mistakes while developing new skills (Sweigart et al., 2014). Moreover, digital interventions are easily scalable, and many are accessible to anyone with an internet connection (Goldberg et al., 2022).

Despite the promise of digitally delivered empathy training, current findings on the efficacy of digital interventions for empathy are mixed. While many studies of digital interventions show encouraging results for improving mental health (Linardon et al., 2024), some indicate that digital platforms can hinder the impact of an empathy intervention. Porcino et al. (2017), for example, found that VR can cause motion sickness, disorientation, and other discomfort that mitigated the intervention's effects. Likewise, some video game and simulation interventions indicate that participants end up increasing empathy only for characters with similar identities or backgrounds in comparison to characters of other cultures and identities (Paiva et al., 2005). Some critics also argue that digital interventions are unlikely to be beneficial, as digital culture has reduced social interaction, increased individualism, and decreased empathic responses among younger generations (Gorry, 2009; Misra et al., 2016).

Present Study

Given the mixed findings on the efficacy of digital interventions for empathy, the present meta-analysis sought to synthesize the evidence from the available randomized controlled trials (RCTs) testing these interventions. We focused specifically on empathy training in adults, given previous research on empathy predominantly focused on early stages of life and less is known about empathy development later in development (Rosenthal & Gatt, 2010; Schonert-Reichl et al., 2012). Demonstrating that digital interventions have the capacity to increase empathy in adulthood may have important implications for individual, relational, and societal health. We hypothesized that digital interventions would have a positive impact on empathy development in adults.

In addition to evaluating the overall impact of digital intervention on empathy in adults, we were also interested in factors that may influence the effectiveness of this approach. We examined the following moderators: neurological or psychological disorder, control group types, career type, type of digital interventions, age, dose, deliberate practice or feedback. Specifically, we hypothesized that digital interventions would have a smaller impact on empathy among individuals with conditions associated with impaired empathic capacity (e.g., autism spectrum disorder, traumatic brain injury, antisocial personality disorder, or schizophrenia). We also expected larger effects when interventions were compared to non-specific controls (i.e., comparisons not intended to be therapeutic) vs. active controls (i.e., those intended to be therapeutic; Wampold et al., 1997), among adults in helping professions vs. other careers, for VR-based interventions compared to other digital formats, for younger vs. older adults, for interventions with higher dosage or longer duration, and when deliberate practice and/or feedback was included.

Methods

Protocol and Registration

The current study was preregistered through the Open Science Framework (https://osf.io/b9ukd?view_only=b70bba762eee4c11ac3528ddd3e79b38) and was conducted based on the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines (Moher et al., 2009).

Eligibility Criteria

The purpose of this meta-analysis was to examine the effectiveness of digital interventions in enhancing empathy in adults. Studies were selected based on the following criteria: a) RCT, b) delivered a digital intervention, c) utilized an empathy outcome measure that included a cognitive, affective, or behavioral component of empathy, or any combination of the three, and d) included an adult population (i.e., samples $\geq$18 years old, on average). We focused on interventions that were designed to be delivered digitally, such as a VR immersive experience or an interactive smartphone app. Examples of interventions that are digital but were not included in this meta-analysis were video games used purely for entertainment purposes, traditional psychotherapy delivered via telehealth, or a psychoeducational document delivered on an e-reader. The first example is not an intervention, and the latter examples involve a digital component, but the digital aspect of the experience is not central to the intervention's delivery nor its effectiveness. Of note, if a study examined cognitive empathy only, they needed to explicitly name it as "cognitive empathy" rather than "theory of mind" or other terms. Studies with non-digital control conditions were eligible. When statistics used to calculate effect sizes were not included in the published findings, the original study authors were contacted to provide necessary data. There were no geographic, cultural, or publication date restrictions for study

inclusion in this meta-analysis. Studies published in English, Spanish, and German were included.

Information Sources

A systematic literature search was conducted on February 23, 2024 to locate relevant studies that met inclusion criteria. The following databases were searched: PsycINFO, PubMed/Medline, CINAHL Plus with Full Text, Cochrane Trials, Scopus, Web of Science, and ProQuest Dissertations and Theses Global. In addition, we utilized the reference lists of previously published meta-analyses (e.g., Teding van Berkhout & Malouff, 2016) to retrieve additional relevant studies. The study was pre-registered prior to conducting the literature search in order to increase the transparency and rigor of the evaluation.

Search

The key terms used for the literature search included: (technology OR digital OR virtual OR internet) AND (intervention OR program OR training OR treatment) AND (mentalizing OR "theory of mind" OR "emotion recognition" OR perspective-taking OR "empath*") AND ("random*"). These search terms were developed by considering the most common words used to describe each factor in our research question, such as terms like "mentalizing," which is synonymous with cognitive empathy (Cerniglia et al., 2019). A research librarian was also consulted for best practices in developing an exhaustive list of search terms. In order to reduce publication bias (Borenstein et al., 2009), dissertations and unpublished findings were included in the literature search. Databases were searched since inception.

Study Selection

The first author and one additional coder independently coded all studies for inclusion/exclusion. Coder training included an overview of the coding manual followed by an

initial phase of coding a small number of titles/abstracts that we then examined together to discuss discrepancies. After all studies were coded by both coders independently, coders reconvened to discuss any discrepancies. Discrepancies were resolved through discussion and consensus, returning to original articles when indicated and consulting with the senior author as needed. Interrater reliabilities for inclusion at title and/or abstract and full text levels were good to excellent (i.e., $ICC \geq 0.60$; Cicchetti, 1994).

Data Collection Process

Standardized spreadsheets were used to record study characteristics and effect size information for each of the trials. Data were independently extracted by the first author and a second coder. Studies were excluded if they were still recruiting participants or if authors did not respond when contacted regarding availability of data.

Data Items

All the studies included in this meta-analysis were intervention studies. Pre- and post-test means and standard deviations from the intervention and control groups were utilized to derive between-group effect sizes. For studies that provided post-test outcome data only (missing baseline data), we compared the groups at post-test. When baseline ns were not reported, we used post-test ns.

Several continuous and categorical variables were coded for descriptive purposes and for moderator testing. Included studies were coded on characteristics of the study and study design, characteristics of the sample, and characteristics of the intervention. Characteristics of the study and study design included: a) publication year, b) method of group assignment, c) number of groups, d) type of control group, and e) empathy measure used. Characteristics of the sample included: a) mean age, b) percentage female, c) percentage racial/ethnic minority, d) diagnosis

(i.e., autism spectrum disorder, schizophrenia, traumatic brain injury, antisocial personality disorder, or none), e) career type (i.e., whether or not the participants had a career in healthcare or education), f) stage of career (e.g., trainee vs non-trainee), and g) country of origin. Characteristics of the intervention included: a) group vs. individual delivery, b) type of digital intervention (i.e., VR, video game, simulation, online module(s), smartphone app, other), c) whether the intervention was delivered in combination with human support (e.g., smartphone application + in-person exercises), d) comparison condition or control group type (i.e., active vs. non-specific; Wampold et al., 1997), f) dosage (i.e., hours), and g) whether the intervention included deliberate practice or feedback.

Summary Measures

As the studies included in this meta-analysis were all RCTs of interventions, effect sizes were calculated as Becker's Δ (1988), defined as the difference between within-group standardized mean differences (i.e., Cohen's *d*). For studies that did not report Cohen's *d,* pre- and post-test means and standard deviations were used to calculate it (Cohen, 1988). Becker's (1988) Δ accounts for potential baseline differences rather than relying exclusively on between-group differences at post-test. The formulas that were utilized to make these calculations with pre- and post-test means and standard deviations are listed below:

$$d_{within} = \frac{M_{pre} - M_{pre}}{SD_{pooled}}$$

(Equation 1)

$$var\ (d_{within}) = (\frac{1}{n} + \frac{d^2}{2n}) \times 2(1 - r)$$

(Equation 2)

11

If the included studies did not report r (the correlation between pre- and post- scores), a correlation of $r_{xx} = .50$ was assumed between time points (Hoyt & Del Re, 2018). It is also important to note that the pooled SD in the denominator of the above equation is from pre-treatment (Becker, 1988).

In order to account for the upward bias of Cohen's d (an overestimation of the absolute value of standardized mean difference in smaller samples), d values were converted to Hedge's g using correction factor J (Borenstein et al., 2009). The between-group effect size for the difference between the digital intervention and control groups was then calculated using Becker's Δ (1988):

$$\Delta = g^T_{within} - g^C_{within}$$

(Equation 3)

$$var(\Delta) = var(g^T_{within}) + var(g^C_{within})$$

(Equation 4)

Superscript T refers to the digital intervention and C refers to comparison condition. The Becker's Δ formula can be reversed, so it is important to note the order (Treatment – Control) for interpretational implications. In this case, a positive value will favor Treatment, and a negative value favors Control. For each reported effect size in each study included in this meta-analysis, calculations were completed using the 'metafor' (Viechtbauer, 2010) and 'MAd' packages in R (Del Re & Hoyt, 2014). A random-effects model was used, with effect sizes weighted based on estimates of both within-study and between-study variance. For studies that measured cognitive and affective empathy outcomes separately, we combined them into a single value. Effects at follow up time points were analyzed separately, when available.

Moderator Analysis

We hypothesized that various factors (e.g., neurological or psychological disorder, career type, type of digital intervention, etc.) may influence the impact of digital interventions on empathy in adults. We examined both continuous (e.g., age of participants) or categorical (e.g., type of profession, type of digital intervention) moderators using meta-regression methods implemented in the 'metafor' package in R (Viechtbauer, 2010). Continuous moderators included age and treatment length (in weeks). Categorical moderators included presence of a neurological or psychological disorder, control group type, career type, deliberate practice or feedback, and type of digital intervention. These moderators were dummy coded and included as predictors in the meta-regression.

Synthesis of Results

Aggregation of Effect Sizes within Studies

As some studies reported multiple effect sizes that corresponded to different measures of empathy, aggregation at the study level was necessary. After calculating Becker's Δ, as described above, all dependent effect sizes within a given study were aggregated using the 'MAd' package (Del Re & Hoyt, 2018), in accordance with Hunter & Schmidt's (2004) recommendations.

Tests of Heterogeneity

The variability of the effect sizes was examined using a heterogeneity test to determine if variability is greater than what would be expected by chance. Specifically, we examined the Q and I^2 statistics. The Q statistic was used to test whether the observed between-study variability was greater than would be expected by chance alone (Borenstein et al., 2009). The I^2 was calculated to determine what percentage of the observed variance reflects actual differences among studies beyond the effects of random error (Borenstein et al., 2009). The extent of

heterogeneity can be interpreted based on Higgins and Thompson's (2002) guidelines: 25-49% is small, 50-74% is medium, and 75-100% is a large amount of heterogeneity. A significant Q or an I^2 in the medium, or large range indicates that variation from the true effect size may be due to (measured or unmeasured) moderating variables.

Assessment of Reporting Bias

Studies with significant results tend to be published more so than studies with less significant results (Borenstein et al., 2009), which ultimately influences results of meta-analyses if not accounted for with assessment of reporting bias. The most effective way to reduce bias is to perform a comprehensive and exhaustive literature search. In addition, we examined the impact of potential publication bias using the trim-and-fill procedure (Duval and Tweedie, 2000). Trim-and-fill analysis may be adequate for assessing publication bias, but if a meta-analysis has high heterogeneity, trim-and-fill's ability to correct for bias may be limited (Carter et al., 2019; Terrin et al., 2003). Moreover, trim-and-fill analyses are based on a fixed effects model, rather than the random effects model we used. Thus, trim-and-fill analyses were interpreted cautiously. Subsequently, visual inspection of a contour-enhanced funnel plot allowed for more thorough assessment of publication bias. Contour lines added to the funnel plot mark conventional milestones in levels of statistical significance and can aid in discerning if studies are missing in areas of statistical non-significance vs. higher significance. Studies missing in areas of statistical non-significance suggests that asymmetry is due to publication bias, whereas studies missing in areas of higher statistical significance may suggest factors other than publication bias, such as variable study quality (Peters et al., 2008).

Results

Study Selection

The search yielded 1521 citations. We removed 644 duplicates and evaluated the remaining 877 studies' titles and abstracts based on the inclusion criteria. This evaluation resulted in 184 studies eligible for full-text review (Figure 1). After screening with inclusion and exclusion criteria, 24 studies were retained, representing 3,137 participants. Two of the citations (Hattink et al., 2015; Silveira et al., 2023) each included two studies within their publication, resulting in a total of 26 comparisons included in the meta-analysis.

Study Characteristics

Descriptive statistics of study and sample characteristics can be viewed in Table 1 (see study and sample characteristics and effect size data for each individual study: https://osf.io/ak6ym/?view_only=f9e2d1466bf74b4c819c2038bb35ae00). The included studies were published between 2007 and 2024. Almost half of the studies took place in North America (45.8%), and most of the remaining studies occurred in Europe (41.7%), with a minority in Asia (12.5%) and one in Australia (4.2%). Participants were on average 35.44 years old ($SD = 12.17$), 73.17% female, 18.73% racial/ethnic minorities, and 84.71% had some post-secondary education. The most common study population consisted of individuals in helping professions (e.g., healthcare professionals; 53.85%). Other samples were recruited from the general population (30.77%), college and graduate students (11.55%), and broader university communities (3.85%). Only one study included in this meta-analysis included a sample population with a clinical diagnosis, which was Parkinson's disease. The majority of studies used non-specific control conditions (61.54%). The digital interventions used included self-guided online modules (30.77%), VR (26.92%), smartphone apps (19.23%), video games (11.54%), and audio and/or video (11.54%). The majority of studies did not include a follow-up assessment

time point (76.92%). For studies with a follow-up assessment, the average follow-up length post-treatment was 2.16 weeks ($SD = 3.92$, range = one to 12 weeks).

Effectiveness of Digital Interventions on Empathy

The overall estimated effect size ($g = 0.19$, 95% CI [0.05, 0.32]) was statistically significant ($p = .006$), indicating that digital interventions have a small positive effect on empathy in adults on average (see Figure 2 for forest plot). However, when accounting for publication bias using trim-and-fill adjustment, the estimated effect size became smaller ($g = 0.07$, 95% CI [-0.07, 0.22]) and non-significant ($p = .322$; see Figure 3 for funnel plot depicting results of trim-and-fill adjustment and Figure 4 for color contoured funnel plot). Specifically, the trim-and-fill analysis indicated an under-representation of studies with larger standard errors (i.e., smaller studies) showing non-significant results, which suggests that the initial positive effect size might have been inflated due to publication bias (i.e., publishing smaller studies with disproportionately positive results). Additionally, heterogeneity was quite high both before (67.84% [95% CI 48.33, 85.80]) and after trim-and-fill analysis (76.06%, [95% CI [63.11, 88.14]), indicating the presence of unexplained variability between studies. Studies with follow-up data were analyzed separately ($k = 6$) and a small effect was found ($g = 0.29$, 95% CI [0.07,0.5], $p = .008$). This effect was robust to trim-and-fill analysis and no missing studies were imputed, though trim-and-fill was likely underpowered due to the small number of studies (Duval & Tweedie, 2000).

Moderator Tests

The moderators evaluated in this meta-analysis included age, diagnosis, career type (e.g., healthcare professional), dose (i.e., treatment length), deliberate practice or feedback, control group type (i.e., active vs. non-specific), and type of digital intervention (e.g., VR). Contrary to

our hypotheses, no significant moderators were identified, (see Tables 2 and 3 for moderator test results).

Results of Additional Analysis

Given that control groups are known to impact effects in psychological interventions generally (Goldberg at al., 2022; Wampold & Imel, 2015) as well as in digital interventions (Goldberg et al., 2023; Linardon et al., 2019), we conducted an exploratory analysis examining effects for studies using non-specific control groups ($k = 16$) and active control groups ($k = 10$) separately. This analysis revealed a small positive effect for studies with non-specific control groups ($k = 16$, $g = 0.22$, 95% CI [0.12, 0.33], $p < .001$), and it was robust to publication bias. After trim-and-fill analysis where one study was imputed below the omnibus effect, the effect remained significant ($g = 0.21$, 95% CI [0.1, 0.32], $p < 0.001$). The effect for studies with active controls was not significant ($g = 0.13$, 95% CI [-0.18, 0.44], $p = .412$) and remained non-significant with trim-and-fill analysis ($g = 0.04$, 95% CI [-0.29, 0.37], $p = .801$).

Discussion

Summary of Findings

The development of digital interventions for improving the psychosocial skill of empathy is a recently emerging area of scientific interest. We examined 24 RCTs (26 comparisons; $n = 3,137$) comparing digital interventions with non-digital control conditions. Results indicated a small, positive effect of digital interventions on enhancing empathy in adults ($g = 0.19$, 95% CI [0.05, 0.32], $p = .006$), though once publication bias was taken into account, this result was rendered non-significant. Interestingly, when examining studies with follow-up data separately ($k = 6$), a small effect was found ($g = 0.29$, 95% CI [0.07, 0.5], $p = .008$) that was robust to publication bias, although trim-and-fill adjustment was likely underpowered. Additionally, a

small effect was found for studies with non-specific control groups ($g = 0.22$, 95% CI [0.12, 0.33], $p < .001$), although this was an exploratory analysis and was not a preregistered primary analysis. These findings raise several questions regarding the relationship between digital interventions and empathy in adults.

Empathy May Require Human-Supported Interventions

The modest results from the current study contrast with a recent meta-analysis demonstrating the effects of various interventions in enhancing empathy, with the majority of studies focusing on traditional, non-digital approaches ($g = 0.63$, $p < .001$; Teding van Berkhout & Malouff, 2016). Our results may indicate that empathy training is not particularly well-supported by digital platforms. Empathy skills may be distinct from the types of skills that are successfully supported by digital platforms, such as mindfulness skills (Linardon et al., 2024), symptom tracking skills for addressing mental and physical health concerns (Serre et al., 2023), and language learning skills (Lee & Borah, 2020). Empathy is a non-self-focused skill that requires understanding and connecting with the emotions of others (Cuff et al., 2016). Given the organic development of empathy that occurs in the context of human relationships (Davidov et al., 2013), it could be unrealistic to expect engaging with a digital platform to be an adequate surrogate for producing empathy.

Another important difference between empathy and other skills that are supported by digital interventions is that empathy may be more emotionally demanding (Casey et al., 2013). Willingness to tap into the perspective and emotional experience of another, particularly if the other is going through something emotionally challenging, requires vulnerability to experiencing distress oneself. In fact, it is often the goal of empathy to join someone in their pain (Jankowiak-Siuda et al., 2011). Other skills such as relaxation skills for anxiety or behavioral activation for

depression do not require taking on others' pain in addition to one's own. Given the emotional labor involved in empathy, emotion regulation skills may be an essential pre-requisite that allows one to manage personal distress while remaining connected to the other (Gerdes et al., 2010; Decety and Jackson, 2004). If an individual lacks emotion regulation skills, empathizing with others can be potentially dangerous, which is one reason many people are instinctively limited in their willingness to empathize (Wood et al., 1990). Moreover, the amount of exposure one has to giving and receiving empathy in relationships throughout one's life is associated with the strength of one's own empathic capacity (Stern et al., 2024). This highlights another distinction between empathy skills and fitness, language, mindfulness, or mental health symptom reduction skills; the latter skills do not require a reservoir of specific emotional experiences from which the skill can grow. Digital interventions, which oftentimes have more modest effects than human-delivered interventions (e.g., Linardon et al., 2024), may simply not be potent enough to compensate for these influential developmental processes. It is possible that the development of empathy may be better supported by human-delivered interventions or a blend of digital interventions with human support, which can provide more personalized guidance, attunement to non-verbal communication (Turkle, 2011), and offer genuine connection.

Technological Limitations May Restrict the Effectiveness of Digital Interventions on Empathy

Conversely, the non-significant trim-and-fill adjusted effect size of the current study might indicate that digital interventions are simply limited by the current technology available. The analysis of studies with follow-up data had a small but significant effect ($g = 0.29$, $p = .008$) as well as an exploratory analysis of studies with non-specific control groups ($g = 0.22$, $p < .001$), both of which were robust to publication bias. While these effects are small, they suggest

that digital interventions may be effective in the longer term and relative to non-specific control groups. Another meta-analysis, which focused on VR interventions specifically (Martingano et al., 2021), compared baseline empathy (pre-test or control condition) to empathy levels following a VR experience and found that VR was effective ($d = 0.44$, 95% CI [0.23,0.66], $p = <.001$). Although we did not find that intervention type moderated effects in the current meta-analysis, Martingano et al.'s findings and the current study's small but significant findings for subsets of studies suggest that some digital interventions may still hold promise for enhancing empathy in adults.

As evidenced by Martingano et al.'s (2021) study, VR may be promising as an approach for enhancing empathy, although it is noteworthy that VR ultimately was no more effective than traditional interventions (Martingano et al., 2021). Research regarding VR's effectiveness for fostering empathy is still in its early stages and findings are mixed overall (Martingano et al., 2021; Porcino et al., 2017), which is also reflected in the lack of a significant effect for VR as a treatment type in the current study's moderator analysis. It may be important to explore alternative answers to the question of what type of technology could be more effective, to replace or pair with VR, for an empathy-focused digital intervention.

One answer proposed by psychotherapy researchers who examine interventions for psychosocial skills training is that interactive capability may be what is missing from current digital interventions (Imel et al., 2019). Machine learning technology that can process participants' language and facial expressions and respond appropriately in real time may hold the most promise for cultivating a deeply human trait and skill such as empathy (Creed et al., 2022; Imel et al., 2019). Machine learning tools for psychotherapy training are in their infancy but thus far have demonstrated an impressive ability to identify meaningful verbalizations relevant to

expressing empathy in the context of counseling (Kuo et al., 2023). There are, of course, rapid advances occurring in this area in recent years through the use of large language models (e.g., Hume AI, 2025). Interactive technology that can detect the subtleties and nuances of human expression may be able to mimic a real interaction and subsequently elicit more genuine empathy.

Another question that the current findings raise is whether digital interventions might be more effective if they incorporated a feedback component. Several of the studies included in this meta-analysis incorporated practice of an interaction (e.g., a medical professional practicing demonstrating empathy in a conversation with a simulated patient; Deladisma et al., 2007; Juan, 2023; Kaltman et al., 2018; Maggio et al., 2024), however, only one study included feedback in response to the practice as part of the intervention model (Kaltman et al., 2018). In psychotherapy research, RCTs of deliberate practice (i.e., focused, goal-oriented skill repetition with feedback) interventions have been shown to improve interpersonal skills (Larsson et al., 2023; Olsson et al., 2024). Additionally, Teding van Berkhout and Malouff (2016) found that use of behavioral skills training principles (instruction, modeling, practice, and feedback) was a moderator with slightly higher effect sizes than other studies. Feedback regarding demonstrations of empathy is likely a critical yet absent component from the interventions in the current meta-analysis. Advances in machine learning and artificial intelligence may well be able to provide the kind of feedback that is helpful for learning empathy.

Limitations of the Current Study

One limitation inherent in the process of conducting a meta-analysis is the limited amount of published literature available to include in our examination. As there were only 24 studies that met inclusion criteria, the moderator analyses may have been underpowered to detect

small effects (Hedges & Pigott, 2004). Additionally, authors did not consistently report potential moderators of interest, including mean age, which limited our ability to test them as predictors of effectiveness of treatment. Some publication bias assessments (e.g., for effects at follow up) were very likely underpowered (Duval & Tweedie, 2000).

Another limitation is that empathy continues to be inconsistently operationalized and measured across studies which makes it difficult to compare findings with confidence. Most of the included studies only used self-report measures to assess how well participants felt they had managed to take on the perspective or the emotions of another, which operationalizes empathy as primarily an internal experience without a behavioral component. Other studies focused more on participants' ability to demonstrate empathy behaviorally, classifying empathy as a skill. Variability in how empathy is conceptualized across intervention studies is a significant limitation to empathy development research.

Finally, self-report measures, which are at higher risk of inaccuracy than objective measures due to response bias and subjectivity, were utilized exclusively by most studies included in this meta-analysis. Teding van Berkhout and Malouff (2016) along with previous, unpublished meta-analyses (Butters, 2010; Reed, 1996), found that changes in empathy tended to be larger when objective measures were used compared to self-report measures. In the current meta-analysis, there were too few studies that utilized an objective measure to conduct an exploratory moderator analysis on this characteristic.

Conclusions and Future Research Directions

The psychosocial trait and skill of empathy is foundational to human relationships and functioning communities (De Paul & Guibert, 2008; Gibb & Abela, 2008). As conflict around the world becomes increasingly pervasive (Armed Conflict Location & Event Data, 2024), there

is arguably a pressing need for developing this capacity at scale. Research on the effectiveness of digital interventions for enhancing empathy in adults is still in its infancy, though the evidence thus far indicates potential for digital interventions to have a modest and potentially non-specific (i.e., not superior to active controls) positive impact. Digital interventions are worth continued investment and investigation given their scalability, accessibility, and cost-effectiveness (Goldberg et al., 2022; Linardon et al., 2019; Linardon et al., 2025). Ideally future studies will investigate machine learning technology that allows for interactive experiences, tracking of verbal and non-verbal communication (e.g., facial expressions), as well as an ability to offer feedback. The most effective yet still scalable intervention for empathy may reveal itself to be a blend of digital and human-delivered support, considering the complex and relationship-based conditions typically required in traditional or organic contexts for empathy development. Next steps for research on digital interventions for enhancing empathy in adults will hopefully involve larger, higher-powered studies, an opportunity for deliberate practice and feedback, a combination of self-report and objective, behavioral measures of empathy, and interactive artificial intelligence components.

References

Armed Conflict Location & Event Data. (2024, January). *Conflict Index Results: January 2024*. https://acleddata.com/conflict-index/index-january-2024/

Becker B (1988). Synthesizing standardized mean-change measures. *British Journal of Mathematical and Statistical Psychology*, 41, 257–278. https://doi.org/10.1111/j.2044-8317.1988.tb00901.x

Boltz, L., Henriksen, D., & Mishra, P. (2015). Rethinking technology & creativity in the 21st

century: empathy through gaming-perspective taking in a complex world. *TechTrends:*

Linking Research & Practice to Improve Learning, *59*(6), 3–8.

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2009). Introduction to

Meta-Analysis. John Wiley & Sons.

Bunn, W., & Terpstra, J. (2009). Cultivating empathy for the mentally ill using simulated

auditory hallucinations. *Academic Psychiatry*, *33*, 457-460.

https://doi.org/10.1176/appi.ap.33.6.457

Butters, R. P. (2010). *A meta-analysis of empathy training programs for client populations*

[Doctoral dissertation, University of Utah].

Carter, E. C., Schönbrodt, F. D., Gervais, W. M., & Hilgard, J. (2019). Correcting for bias in

psychology: A comparison of meta-analytic methods. *Advances in Methods and*

Practices in Psychological Science, *2*(2), 115-144.

https://doi.org/10.1177/2515245919847196

Casey, H., Rogers, R. D., Burns, T., & Yiend, J. (2013). Emotion regulation in

psychopathy. *Biological Psychology*, *92*(3), 541-548.

https://doi.org/10.1016/j.biopsycho.2012.06.011

Cerniglia, L., Bartolomeo, L., Capobianco, M., Lo Russo, S. L. M., Festucci, F., Tambelli, R., . .

., Cimino, S. (2019). Intersections and Divergences Between Empathizing and

Mentalizing: Development, Recent Advancements by Neuroimaging and the Future of

Animal Modeling. *Frontiers in Behavioral Neuroscience, 13*, 212.

https://doi.org/10.3389/fnbeh.2019.00212

Chen, H. C., Ignacio, J., & Yobas, P. (2020). Evaluation of the symptom-focused health

    assessment and empathy program for undergraduate nursing students: A randomized

    controlled trial. *Nurse Education Today*, *94*, 104566.

    https://doi.org/10.1016/j.nedt.2020.104566

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Lawrence

    Earlbaum Associates.

Creed, T. A., Salama, L., Slevin, R., Tanana, M., Imel, Z., Narayanan, S., & Atkins, D. C.

    (2022). Enhancing the quality of cognitive behavioral therapy in community mental

    health through artificial intelligence generated fidelity feedback (Project AFFECT): a

    study protocol. *BMC Health Services Research*, *22*(1), 1177.

    https://doi.org/10.1186/s12913-022-08519-9

Cuff, B. M. P., Brown, S. J., Taylor, L., & Howat, D. J. (2016). Empathy: A Review of the

    Concept. *Emotion Review*, *8*(2), 144–153. https://doi.org/10.1177/1754073914558466

Davidov, M., Zahn-Waxler, C., Roth-Hanania, R., & Knafo, A. (2013). Concern for others in the

    first year of life: Theory, evidence, and avenues for research. *Child Development*

    Perspectives, 7, 126–131. https://doi.org/10.1111/cdep.12028

Deladisma, A. M., Cohen, M., Stevens, A., Wagner, P., Lok, B., Bernard, T., ... & Lind, D. S.

    (2007). Do medical students respond empathetically to a virtual patient?. *The American*

    Journal of Surgery, *193*(6), 756-760. https://doi.org/10.1016/j.amjsurg.2007.01.021

Delaney, R., Nelson, C., Pacifici, C., White, L., & Smalley, B. K. (2012). Web-enhanced

    preservice training for prospective resource parents: A randomized trial of effectiveness

    and user satisfaction. *Journal of Social Service Research*, *38*(4), 503-514.

    https://doi.org/10.1080/01488376.2012.696416

Decety, J., & Jackson, P. L. (2004). The functional architecture of human empathy. *Behavioral and Cognitive Neuroscience Reviews, 3*(2), 71-100. https://doi.org/10.1177/1534582304267187

Decety, J., & Meyer, M. (2008). From emotion resonance to empathic understanding: A social developmental neuroscience account. *Development and Psychopathology*, *20*(4), 1053-1080. https://doi.org/10.1017/S0954579408000503

Del Re, A. C. & Hoyt, W. T. (2014). MAd: Meta-analysis with mean differences. R package version 0.8-2.

De Paul, J., & Guibert, M. (2008). Empathy and child neglect: A theoretical model. *Child Abuse & Neglect*, *32*(11), 1063–1071. https://doi.org/10.1016/j.chiabu.2008.03.003

Duval, S., & Tweedie, R. (2000). Trim and Fill: A Simple Funnel-Plot–Based Method of Testing and Adjusting for Publication Bias in Meta-Analysis. *Biometrics*, *56*(2), 455–463. https://doi.org/10.1111/j.0006-341X.2000.00455.x

Enke, B. (2023). Market exposure and human morality. *Nature Human Behaviour*, *7*(1), 134–141. https://doi.org/10.1038/s41562-022-01480-x

Gerdes, K. E., Segal, E. A., & Lietz, C. A. (2010). Conceptualising and measuring empathy. *British Journal of Social Work, 40*(7), 2326–2343. https://doi.org/10.1093/bjsw/bcq048

Gibb, B. E., & Abela, J. R. Z. (2008). Emotional Abuse, Verbal Victimization, and the Development of Children's Negative Inferential Styles and Depressive Symptoms. *Cognitive Therapy and Research*, *32*(2), 161–176. https://doi.org/10.1007/s10608-006-9106-x

Goldberg, S. B., Lam, S. U., Simonsson, O., Torous, J., & Sun, S. (2022). Mobile phone-based interventions for mental health: A systematic meta-review of 14 meta-analyses of

randomized controlled trials. *PLOS Digital Health*, *1*(1), e0000002.

https://doi.org/10.1371/journal.pdig.0000002

Goldberg, S. B., Sun, S., Carlbring, P., & Torous, J. (2023). Selecting and describing control

conditions in mobile health randomized controlled trials: a proposed typology. *npj Digital Medicine, 6*(1), 181. https://doi.org/10.1038/s41746-023-00923-7

Gorry, A. (2009). Empathy in the virtual world. *The Chronicle of Higher Education, 56*, B10–

B12.

Gutierrez, B., Kaatz, A., Chu, S., Ramirez, D., Samson-Samuel, C., & Carnes, M. (2014). "Fair

Play": a videogame designed to address implicit race bias through active perspective

taking. *Games for Health Journal*, *3*(6), 371-378. https://doi.org/10.1089/g4h.2013.0071

Hattink, B., Meiland, F., van der Roest, H., Kevern, P., Abiuso, F., Bengtsson, J., ... & Dröes, R.

M. (2015). Web-based STAR E-learning course increases empathy and understanding in

dementia caregivers: results from a randomized controlled trial in the Netherlands and the

United Kingdom. *Journal of Medical Internet Research*, *17*(10), e241.

https://doi.org/10.2196/jmir.4025

Hedges, L. V., & Pigott, T. D. (2004). The power of statistical tests for moderators in meta-

analysis. *Psychological Methods, 9*(4), 426. https://doi.org/10.1037/1082-989X.9.4.426

Herrera, F., Bailenson, J., Weisz, E., Ogle, E., & Zaki, J. (2018). Building long-term empathy: A

large-scale comparison of traditional and virtual reality perspective-taking. *PloS One*, *13*(10), e0204494. https://doi.org/10.1371/journal.pone.0204494

Higgins, J. P., & Thompson, S. G. (2002). Quantifying heterogeneity in a meta-

analysis. *Statistics in Medicine*, *21*(11), 1539-1558. https://doi.org/10.1002/sim.1186

Hojat, M., Louis, D. Z., Markham, F. W., Wender, R., Rabinowitz, C., & Gonnella, J. S. (2011). Physician'' empathy and clinical outcomes for diabetic patients. *Academic Medicine*, *86*(3), 359-364. https://doi.org/10.1097/ACM.0b013e3182086fe1

Hoyt, W. T., & Del Re, A. C. (2018). Effect size calculation in meta-analyses of psychotherapy outcome research. *Psychotherapy Research*, *28*(3), 379-388. https://doi.org/10.1080/10503307.2017.1405171

Hume AI. (2025). *Hume AI*. https://www.hume.ai/

Hunter, J. E. & Schmidt, F. L. (2004). *Methods of Meta-Analysis: Correcting Error and Bias in Research Finding* (2nd ed.). Sage Publications.

Imel, Z. E., Pace, B. T., Soma, C. S., Tanana, M., Hirsch, T., Gibson, J., ... & Atkins, D. C. (2019). Design feasibility of an automated, machine-learning based feedback system for motivational interviewing. *Psychotherapy*, *56*(2), 318. https://doi.org/10.1037/pst0000221

Irvine, A., Billow, M. B., Eberhage, M. G., Seeley, J. R., McMahon, E., & Bourgeois, M. (2012). Mental illness training for licensed staff in long-term care. *Issues in Mental Health Nursing*, *33*(3), 181-194. https://doi.org/10.3109/01612840.2011.639482

Jankowiak-Siuda, K., Rymarczyk, K., & Grabowska, A. (2011). How we empathize with others: a neurobiological perspective. *Medical Science Monitor, 17*(1), RA18. https://doi.org/10.12659/MSM.881324

Jewell, D. E. (2020). *The Use of Personal Testimonials to Enhance Gender Bias Literacy in STEM* [Master's Thesis, Purdue University].

Juan, S. (2023). *The Application of Virtual Simulation to Promoting Empathy in Undergraduate Nursing Students* [Doctoral Dissertation, Columbia University].

Kaltman, S., Talisman, N., Pennestri, S., Syverson, E., Arthur, P., & Vovides, Y. (2018). Using

    technology to enhance teaching of patient-centered interviewing for early medical

    students. *Simulation in Healthcare*, *13*(3), 188-194.

    https://doi.org/10.1097/SIH.0000000000000304

Klein, K. J., & Hodges, S. D. (2001). Gender differences, motivation, and empathic accuracy:

    When it pays to understand. *Personality and Social Psychology Bulletin, 27*(6), 720-730.

    https://doi.org/10.1177/0146167201276007

Krznaric, R. (2014). *Empathy: A handbook for revolution*. Random House.

Kuo, P. B., Tanana, M. J., Goldberg, S. B., Caperton, D. D., Narayanan, S., Atkins, D. C., &

    Imel, Z. E. (2023). Machine-Learning-Based Prediction of Client Distress From Session

    Recordings. C*linical Psychological Science, 12*(3), 435-446.

    https://doi.org/10.1177/21677026231172694

Lam, T. C. M., Kolomitro, K., & Alamparambil, F. C. (2011). Empathy training: Methods,

    evaluation practices, and validity. *Journal of Multidisciplinary Evaluation, 7,* 162–200.

    https://doi.org/10.56645/jmde.v7i16.314

Larsson, J., Werthén, D., Carlsson, J., Salim, O., Davidsson, E., Vaz, A., Sousa, D., & Norberg,

    J. (2023). Does deliberate practice surpass didactic training in learning empathy skills? –

    A randomized controlled study. *Nordic Psychology. 77*(1), 39-52.

    https://doi.org/10.1080/19012276.2023.2247572

Lee, D. K. L., & Borah, P. (2020). Self-presentation on Instagram and friendship development

    among young adults: A moderated mediation model of media richness, perceived

    functionality, and openness. *Computers in Human Behavior*, *103*, 57-66.

    https://doi.org/10.1016/j.chb.2019.09.017

Linardon, J., Cuijpers, P., Carlbring, P., Messer, M., & Fuller-Tyszkiewicz, M. (2019). The
efficacy of app-supported smartphone interventions for mental health problems: A meta-
analysis of randomized controlled trials. *World Psychiatry*, *18*(3), 325-336.
https://doi.org/10.1002/wps.20673

Linardon, J., Torous, J., Firth, J., Cuijpers, P., Messer, M., & Fuller-Tyszkiewicz, M. (2024).
Current evidence on the efficacy of mental health smartphone apps for symptoms of
depression and anxiety. A meta-analysis of 176 randomized controlled trials. *World
Psychiatry, 23*(1), 139-149. https://doi.org/10.1002/wps.21183

Linardon, J., Xie, Q., Swords, C., Torous, J., Sun, S., & Goldberg, S. B. (2025). Methodological
quality in randomised clinical trials of mental health apps: systematic review and
longitudinal analysis. *BMJ Mental Health, 28*, e301595.
https://doi.org/10.1136/bmjment-2025-301595

Maggio, M. G., Luca, A., Cicero, C. E., Calabrò, R. S., Drago, F., Zappia, M., & Nicoletti, A.
(2024). Effectiveness of telerehabilitation plus virtual reality (Tele-RV) in cognitive e
social functioning: A randomized clinical study on Parkinson's disease. *Parkinsonism &
Related Disorders*, *119*, 105970. https://doi.org/10.1016/j.parkreldis.2023.105970

Martingano, A. J., Hererra, F., & Konrath, S. (2021). Virtual reality improves emotional but not
cognitive empathy: A meta-analysis. *Technology, Mind, and Behavior, 2*(1).
https://doi.org/10.1037/tmb0000034

Misra, S., Cheng, L., Genevie, J., & Yuan, M. (2016). The iPhone effect: The quality of in-
person social interactions in the presence of mobile devices. *Environment and
Behavior*, *48*(2), 275-298. https://doi.org/10.1177/0013916514539755

Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., & PRISMA Group*, T. (2009). Preferred

reporting items for systematic reviews and meta-analyses: the PRISMA

statement. *Annals of Internal Medicine*, *151*(4), 264-269.

https://doi.org/10.1371/journal.pmed.1000097

Morelli, S. A., Lieberman, M. D., & Zaki, J. (2015). The emerging study of positive

empathy. *Social and Personality Psychology Compass*, *9*(2), 57-68.

https://doi.org/10.1111/spc3.12157

Mueller, K., Prins, R., & de Heer, H. (2017). An online intervention increases empathy,

resilience, and work engagement among physical therapy students. *Journal of Allied

Health*, *47*(3), 196-203.

Nation, J. A., Wertheim, E. H., & Worthington Jr, E. L. (2017). Evaluation of an online self-help

version of the REACH forgiveness program: Outcomes and predictors of persistence in a

community sample. *Journal of Clinical Psychology*, *74*(6), 819-838.

https://doi.org/10.1002/jclp.22557

Olsson, M., Gustafsson, M., Carlsson, J., Rousmaniere, T., Bergbom, S., & Norberg, J. (2024). A

pilot study of a deliberate practice intervention for therapist trainees. *Nordic Psychology*.

165-184. https://doi.org/10.1080/19012276.2024.2313532

Orosa-Duarte, Á., Mediavilla, R., Muñoz-Sanjose, A., Palao, Á., Garde, J., López-Herrero, V., ...

& Rodríguez-Vega, B. (2021). Mindfulness-based mobile app reduces anxiety and

increases self-compassion in healthcare students: a randomised controlled trial. *Medical

Teacher*, *43*(6), 686-693. https://doi.org/10.1080/0142159X.2021.1887835

Overgaauw, S., Güroğlu, B., Rieffe, C., & Crone, E. A. (2014). Behavior and neural correlates of

empathy in adolescents. *Developmental Neuroscience*, *36*(3-4), 210-219.

https://doi.org/10.1159/000363318

Paiva, A., Dias, J., Sobral, D., Aylett, R., Woods, S., Hall, L., & Zoll, C. (2005). Learning by feeling: Evoking empathy with synthetic characters. *Applied Artificial Intelligence*, *19*(3-4), 235-266. https://doi.org/10.1080/08839510590910165

Pang, C., Li, W., Zhou, Y., Gao, T., & Han, S. (2023). Are women more empathetic than men? Questionnaire and EEG estimations of sex/gender differences in empathic ability. *Social Cognitive and Affective Neuroscience*, *18*(1), nsad008. https://doi.org/10.1093/scan/nsad008

Panto, F., Saito, T., Morita, N., & Ogai, Y. (2022). Mental health care for young people using video games: a pilot RCT on the development of a new intervention method toward hikikomori and futoko. *F1000Research*, *11*, 574. https://doi.org/10.12688/f1000research.119764.1

*Penn, D. L., Ivory, J. D., & Judge, A. (2010). The virtual doppelganger: Effects of a virtual reality simulator on perceptions of schizophrenia. *The Journal of Nervous and Mental Disease*, *198*(6), 437-443. https://doi.org/10.1097/NMD.0b013e3181e07d66 *(previously included Kalyanaraman as first author)

Peters, J. L., Sutton, A. J., Jones, D. R., Abrams, K. R., & Rushton, L. (2008). Contour-enhanced meta-analysis funnel plots help distinguish publication bias from other causes of asymmetry. *Journal of Clinical Epidemiology, 61*(10), 991-996. https://doi.org/10.1016/j.jclinepi.2007.11.010

Porcino, T. M., Clua, E., Trevisan, D., Vasconcelos, C. N., & Valente, L. (2017, April). Minimizing cyber sickness in head mounted display systems: design guidelines and applications. In *2017 IEEE 5th International Conference on Serious Games and*

Applications for Health (SeGAH) (pp. 1-6). IEEE.

> https://doi.org/10.1109/SeGAH.2017.7939283

Quail, M., Brundage, S. B., Spitalnick, J., Allen, P. J., & Beilby, J. (2016). Student self-reported

> communication skills, knowledge and confidence across standardised patient, virtual and

> traditional clinical learning environments. *BMC Medical Education*, *16*, 1-12.

> https://doi.org/10.1186/s12909-016-0577-5

Reed, D. L. (1996). *Effects of a combined didactic and experiential death education/empathy*

> *training program on death anxiety and empathic ability of medical students* [Doctoral

> dissertation].

Rigg, E., & van der Wal-Maris, S. (2020). Student teachers' learning about social

> entrepreneurship education–a Dutch pilot study in primary teacher education. *Discourse*

> *and Communication for Sustainable Education*, *11*(1), 41-52.

Rosenthal, M. K., & Gatt, L. (2010). 'Learning to Live Together': training early childhood

> educators to promote socio-emotional competence of toddlers and pre-school

> children. *European Early Childhood Education Research Journal*, *18*(3), 373-390.

> https://doi.org/10.1080/1350293X.2010.500076

Salcido-Cibrián, L. J., Ramos, N. S., Jiménez, Ó., & Blanca, M. J. (2019). Mindfulness to

> regulate emotions: The Mindfulness and Emotional Intelligence Program (PINEP) and its

> adaptation to a virtual learning platform. *Complementary Therapies in Clinical*

> *Practice*, *36*, 176-180. https://doi.org/10.1016/j.ctcp.2019.07.003

Schonert-Reichl, K. A., Smith, V., Zaidman-Zait, A., & Hertzman, C. (2012). Promoting

> children's prosocial behaviors in school: Impact of the "Roots of Empathy" program on

the social and emotional competence of school-aged children. *School Mental Health*, *4*,
1-21. https://doi.org/10.1007/s12310-011-9064-7

Serre, F., Moriceau, S., Donnadieu, L., Forcier, C., Garnier, H., Alexandre, J. M., ... & Craving-
Manager RCT investigator group. (2023). The Craving-Manager smartphone app
designed to diagnose substance use/addictive disorders, and manage craving and
individual predictors of relapse: a study protocol for a multicenter randomized controlled
trial. *Frontiers in Psychiatry, 14*, 1143167. https://doi.org/10.3389/fpsyt.2023.1143167

Silveira, S., Godara, M., & Singer, T. (2023). Boosting empathy and compassion through
mindfulness-based and socioemotional dyadic practice: randomized controlled trial with
app-delivered trainings. *Journal of Medical Internet Research*, *25*, e45027.
https://doi.org/10.2196/45027

Sterkenburg, P. S., & Vacaru, V. S. (2018). The effectiveness of a serious game to enhance
empathy for care workers for people with disabilities: a parallel randomized controlled
trial. *Disability and Health Journal*, *11*(4), 576-582.
https://doi.org/10.1016/j.dhjo.2018.03.003

Stern, J. A., Bailey, N. A., Costello, M. A., Hellwig, A. F., Mitchell, J., & Allen, J.
P. (2024). Empathy across three generations: From maternal and peer support in
adolescence to adult parenting and child outcomes. *Child Development*, 95, 1628–1640.
https://doi.org/10.1111/cdev.14109

Sundar, S. S., Kang, J., & Oprean, D. (2017). Being there in the midst of the story: How
immersive journalism affects our perceptions and cognitions. *Cyberpsychology,
Behavior, and Social Networking*, *20*(11), 672-682.
https://doi.org/10.1089/cyber.2017.0271

Sung, H. C., Su, H. F., Lee, W. L., Yamakawa, M., & Wang, H. M. (2022). Effects of a dementia virtual reality-based training with peer support for home care workers: A cluster randomized controlled trial. *International Journal of Geriatric Psychiatry*, *37*(9). https://doi.org/10.1002/gps.5799

Sweigart, L., Burden, M., Carlton, K. H., & Fillwalk, J. (2014). Virtual simulations across curriculum prepare nursing students for patient interviews. *Clinical Simulation in Nursing*, *10*(3), e139-e145. https://doi.org/10.1016/j.ecns.2013.10.003

Teding van Berkhout, E., & Malouff, J. M. (2016). The efficacy of empathy training: A meta-analysis of randomized controlled trials. *Journal of Counseling Psychology*, *63*(1), 32. https://doi.org/10.1037/cou0000093

Terrin, N., Schmid, C. H., Lau, J., & Olkin, I. (2003). Adjusting for publication bias in the presence of heterogeneity. *Statistics in Medicine*, *22*(13), 2113-2126. https://doi.org/10.1002/sim.1461

Turkle, S. (2011). *Alone together: Why we expect more from technology and less from each other*. Basic Books.

van Wingerden, E., Wouda, M., & Sterkenburg, P. (2019). Effectiveness of m-learning HiSense APP-ID in enhancing knowledge, empathy, and self-efficacy in caregivers of persons with intellectual disabilities: a randomized controlled trial. *Health and Technology*, *9*(5), 893-901. https://doi.org/10.1007/s12553-019-00361-0

Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor package. *Journal of Statistical Software, 36*(3), 1–49. https://doi.org/10.18637/jss.v036.i03

Vinayak, S., & Judge, J. (2018). Resilience and empathy as predictors of psychological wellbeing among adolescents. *International Journal of Health Sciences and Research*, *8*(4), 192-200.

Wampold, B. E., Mondin, G. W., Moody, M., Stich, F., Benson, K., & Ahn, H. N. (1997). A meta-analysis of outcome studies comparing bona fide psychotherapies: Empirically, "all must have prizes". *Psychological Bulletin*, *122*(3), 203–215. https://doi.org/10.1037/0033-2909.122.3.203

Wampold, B. E., & Imel, Z. E. (2015). *The great psychotherapy debate: The evidence for what makes psychotherapy work*. Routledge.

Watson, J. C., Steckley, P. L., & McMullen, E. J. (2014). The role of empathy in promoting change. *Psychotherapy Research*, *24*(3), 286-298. https://doi.org/10.1080/10503307.2013.802823

Wood, J. V., Saltzberg, J. A., Neale, J. M., Stone, A. A., & Rachmiel, T. B. (1990). Self-focused attention, coping responses, and distressed mood in everyday life. *Journal of Personality and Social Psychology*, *58*(6), 1027-1036. https://doi.org/10.1037/0022-3514.58.6.1027

Yu, J., Parsons, G. S., Lancastle, D., Tonkin, E. T., & Ganesh, S. (2020). "Walking in Their Shoes": The effects of an immersive digital story intervention on empathy in nursing students. *Nursing Open*, *8*(5), 2813-2823. https://doi.org/10.1002/nop2.860

Table 1

Descriptive Statistics of Sample and Study Characteristics

Characteristics	k (% of studies)	Mean (SD)
Age		35.44 (12.17)
% Female		73.17 (19.9)
% REM		18.73 (20.72)
% Some College		84.71 (27.39)
Country		
Australia	2 (7.69%)	
Canada	1 (3.85%)	
Germany	2 (7.69%)	
Italy	1 (3.85%)	
Japan	1 (3.85%)	
Netherlands	2 (7.69%)	
Singapore	1 (3.85%)	
Spain	2 (7.69%)	
Taiwan	1 (3.85%)	
UK	1 (3.85%)	
UK and Netherlands	2 (7.69%)	
USA	10 (38.46%)	
Diagnosis		
None	25 (96.15%)	
Parkinson's Disease	1 (3.85%)	
Career Type		
College Students	2 (7.69%)	
General Population	8 (3.08%)	
Graduate Students	1 (3.85%)	
Helper	14 (53.85%)	
University Community	1 (3.85%)	
Deliberate Practice or Feedback		
Yes	7 (26.92%)	
No	19 (73.08%)	
Control Group Type		
Non-specific	16 (61.54%)	
Active	10 (38.46%)	
Treatment Type		
Smartphone App	5 (19.23%)	
Audio and/or Video	3 (11.54%)	
Self-guided Online Modules	8 (30.77%)	
Video Game	3 (11.54%)	
Virtual Reality	7 (26.92%)	

Human-supported Intervention		
Yes	1 (3.85%)	
No	25 (96.15%)	
Treatment Weeks		4.53 (5.53)
Follow-up Length (Weeks)		2.16 (3.92)
Intention-to-treat N		146.96 (99.78)

Note. % REM = Percentage of racial or ethnic minorities.

Table 2

Continuous Moderator Test Results

Predictor	k (# of Studies)	Effect Size	Standard Error	p	95% CI	I² [95% CI]	Q Test for Heterogeneity of Effect Sizes (QE [df] p)	Q Test of Moderators (QM [df] p)
Age	21	0.01	0.01	0.610	-0.01, 0.02	66.41% [41.46, 86.26]	53.53 (df = 19), p < .001	0.27 (df = 1), p = .611
Dose (Treatment Weeks)	25	0.01	0.01	0.230	-0.01, 0.04	69.69% [49.54, 86.51]	68.52 (df = 23), p < .001	1.18 (df = 1), p = .278

Note. $k =$ number of studies; $p =$ p-value; CI = confidence interval; df = degrees of freedom.

Table 3

Categorical Moderator Test Results

Predictor	k (Number of Studies)	Effect Size	Standard Error	p	CI	I² [95% CI]	Q Test for Heterogeneity of Effect Sizes (QE [df] p)	Q Test of Moderators (QM [df] p)
Disorder (Parkinson's disease)	26					69.01% [50.04, 86.68]	69.44 (df = 24), p < .001	0.43 (df = 1), *p* = .510
None (intercept)	25 (96.15%)	0.18	0.07	0.010	0.04, 0.31			
Parkinson's Disease	1 (3.85%)	0.28	0.43	0.510	-0.56, 1.11			
Career Type (all categories)	26					57.53% [32.83, 85.40]	47 (df = 21), *p* = 0.0009	9.13 (df = 4), *p* = .060
Helper (intercept)	14 (53.85%)	0.24	0.08	0.003				
College Students	2 (7.69%)	0.04	0.25	0.861	-0.45, 0.53			
General Population	8 (30.77%)	-0.01	0.14	0.918	-0.28, 0.25			
Grad Students	1 (3.85%)	-0.78	0.3	0.009	-1.36, -0.19			
University Community	1 (3.85%)	-0.45	0.29	0.118	-1.02, 0.12			
Career Type (helper vs. other)	26					67.86% [47.94, 86.14]	66.59 (df = 24), *p* < .001	0.8 (df = 1), *p* = .370
Other (Intercept)	12 (46.15%)	0.12	0.1	0.246	-0.08, 0.32			
Helper	14 (53.85%)	0.12	0.14	0.372	-0.15, 0.39			
Treatment Type	26					67.74% [44.86, 85.81]	61.04 (df = 21), *p* < .001	4.40 (df = 4), *p* = .350
Smartphone App (intercept)	5 (19.23%)	0.22	0.16	0.172	-0.09, 0.53			
Audio and/or Video	3 (11.54%)	0.02	0.25	0.938	-0.46, 0.50			

	k (%)							
Self-guided Online Modules	8 (30.77%)	0.019	0.2	0.927	-0.38, 0.42			
Video Game	3 (11.54%)	-0.43	0.26	0.096	-0.93, 0.08			
Virtual Reality	7 (26.92%)	0.04	0.21	0.856	-0.37, 0.45			
Control Group Type	26					68.15% [48.48, 86.25]	68.01 (df = 24), p <.001	0.56 (df = 1), p = .451
Non-specific Control (intercept)	16 (61.54%)	0.23	0.09	0.010	0.06, 0.40			
Specific Active Control	10 (38.46%)	-0.11	0.14	0.450	-0.38, 0.17			
Deliberate Practice or Feedback	26					67.86% [47.49, 85.73]	68.29 (df = 24), p < .001	1.25 (df = 1), p = .264
None (intercept)	19 (73.08%)	0.23	0.08	0.004	0.08, 0.39			
Deliberate Practice or Feedback	7 (26.92%)	-0.17	0.15	0.264	-0.47, 0.13			

Note. k = number of studies; p = p-value; CI = confidence interval; df = degrees of freedom.

Figure 1

PRISMA Flow Diagram

1521 studies retrieved from the following databases:	
501 PsycINFO, CINAHL Plus with Full Text, and MEDLINE	
385 Web of Science	→ **644** duplicates removed
280 Scopus	
34 ProQuest Dissertations & Theses Global	
284 Cochrane Trials	
37 Hand searching	

877 records screened (title & abstract) → Records excluded (title & abstract):

99 not randomized clinical trial
74 not digital intervention
247 not adults
259 no empathy outcome measure
12 brain imaging study
2 full text not available

184 Reports assessed for eligibility (full text) → Reports excluded (full text):
23 not randomized clinical trial
15 not digital intervention
13 no non-digital control group
45 not adults
53 no empathy outcome measure
4 duplicates
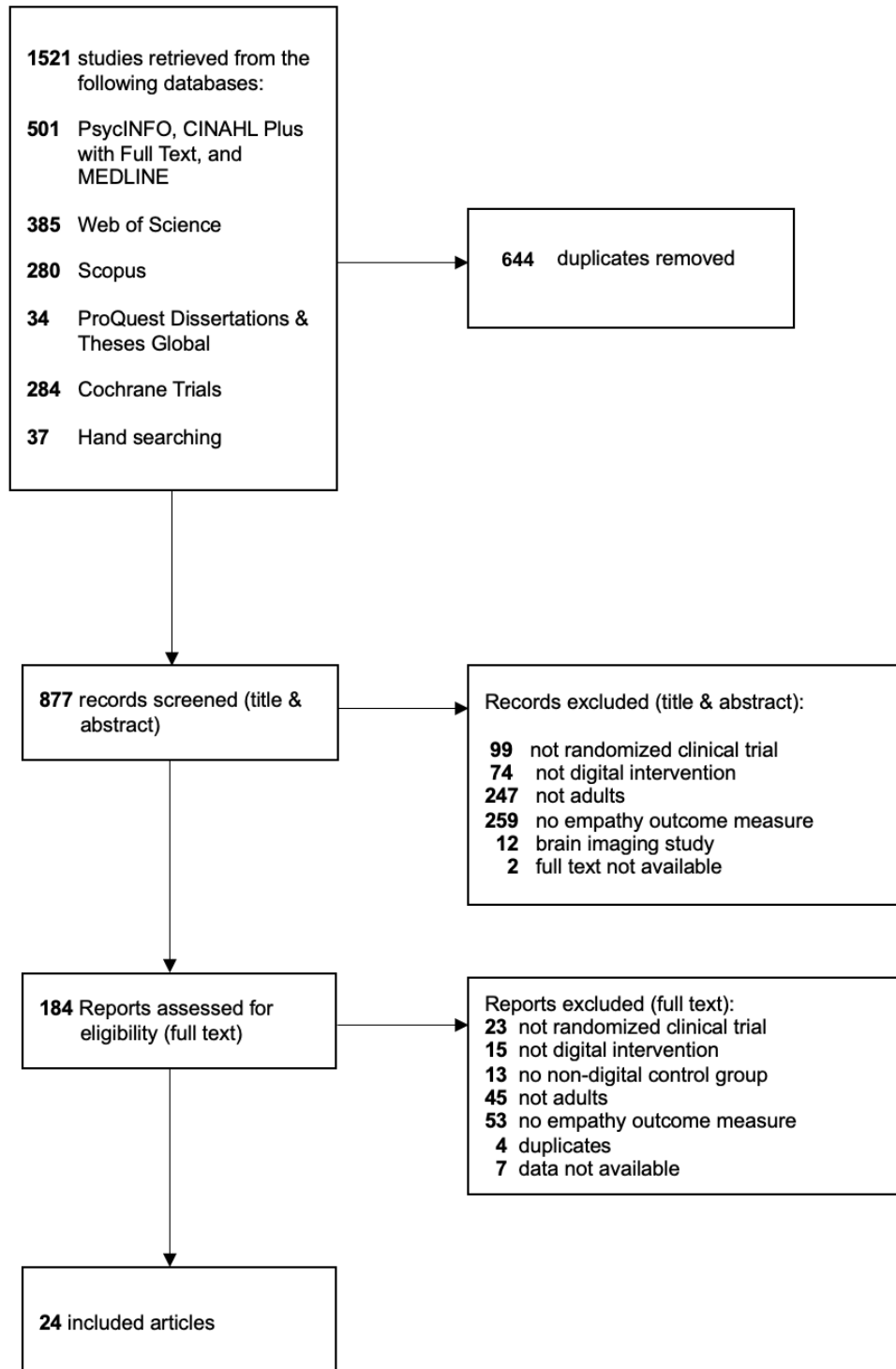**7** data not available

24 included articles

Figure 2

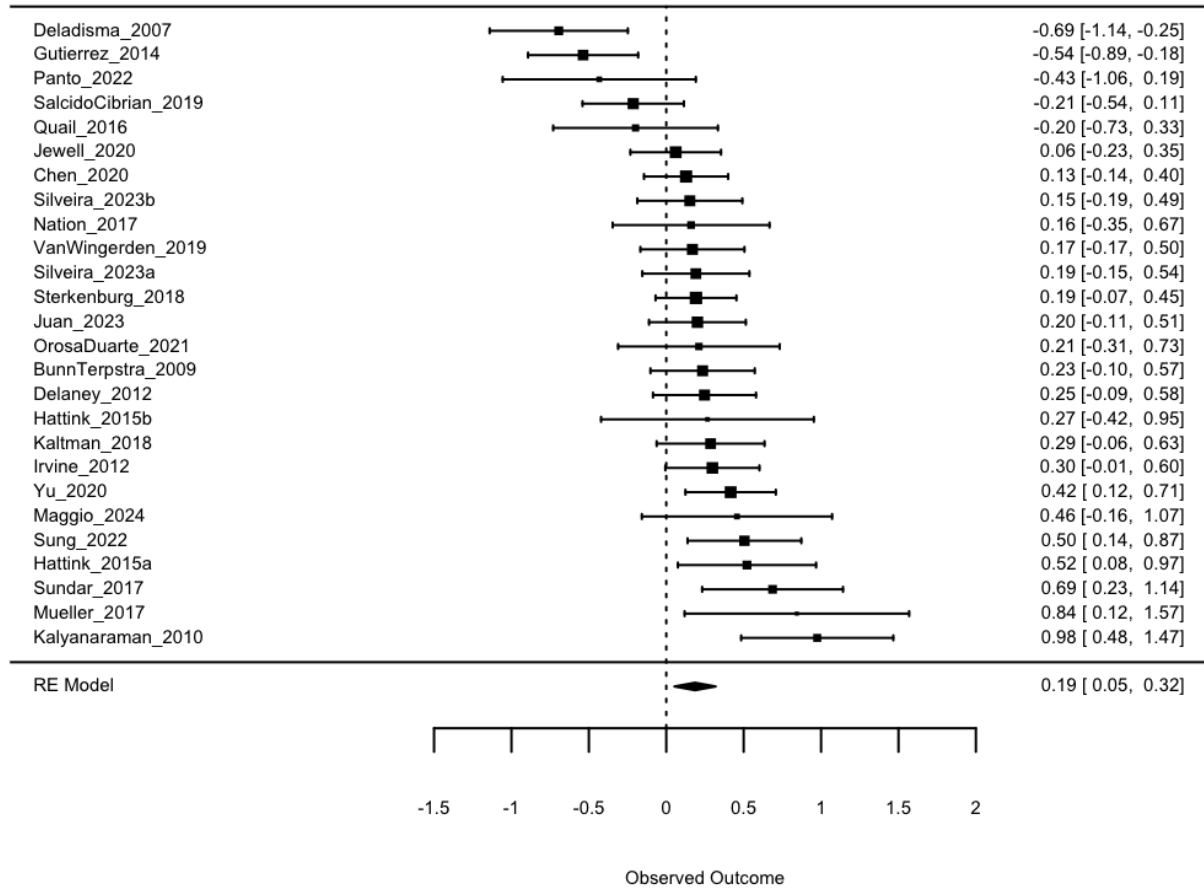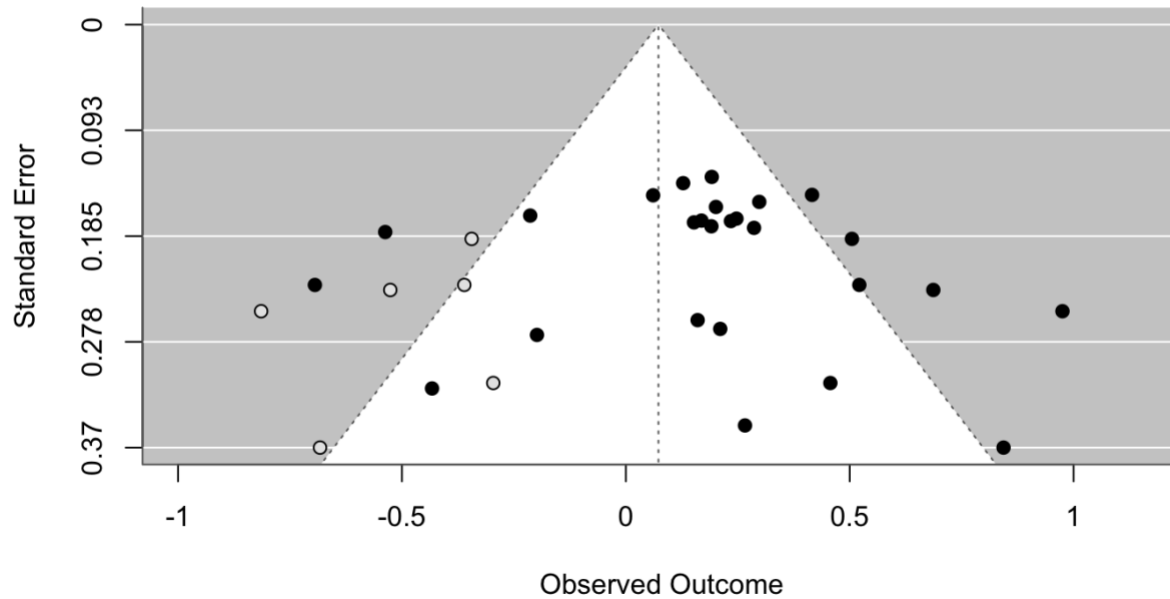*Forest Plot of Effect Sizes Across 26 Comparisons*

Figure 3

Funnel Plot Depicting Results of Trim-and-Fill Adjustment

Note. Funnel plot asymmetry was detected, with smaller standard errors (higher precision) clustered towards the right side of the plot, suggesting a possible bias towards studies with statistically significant positive results. Six studies were imputed in the lower left quadrant based on estimated publication bias.

Figure 4

Color-Coded Enhancement Of Funnel Plot
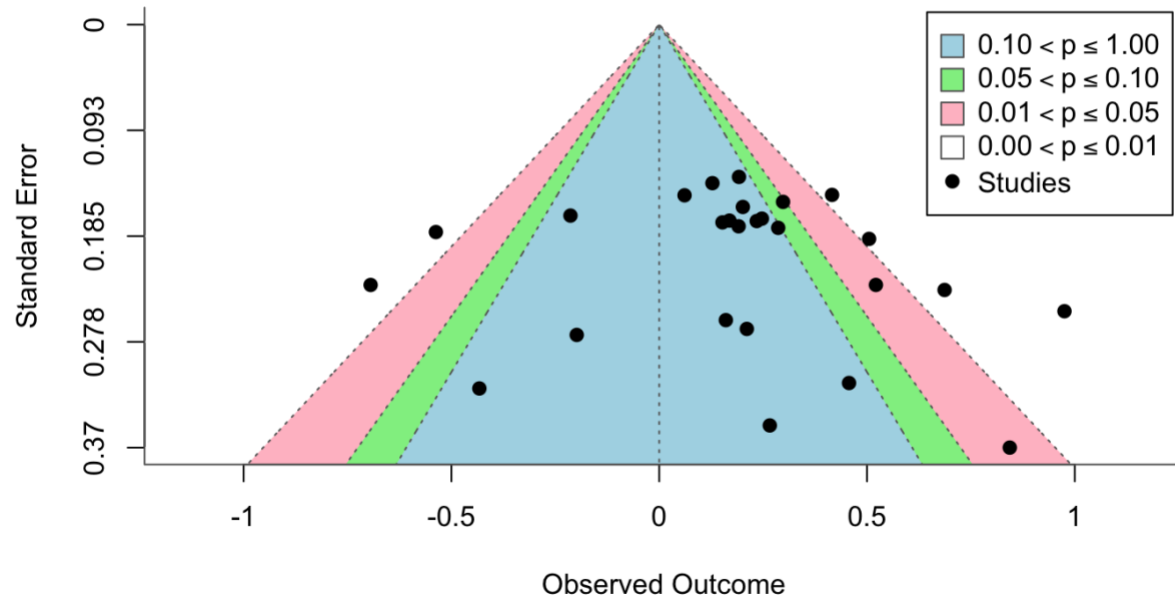
Note. The color contoured plot further aids in detecting potential bias. Some studies, but not a majority, appear in the pink and white regions (areas of higher statistical significance) on the right side of the figure, supporting the possibility of publication bias.